



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 184078

TO: Manjunath N Rao
Location: rem/2A01/2C70
Art Unit: 1652
Tuesday, ~~April 11, 2006~~
Case Serial Number: 10/784870

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:32:32 ; Search time 138.392 Seconds

(without alignments)
2028.756 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059
Sequence: 1 MRKKVPLSLVSAAILSTV.....EYQANVPVGPQPSLAIYN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.21:*
1: geneeqp1980s:*
2: geneeqp1990s:*
3: geneeqp2000s:*
4: geneeqp2001s:*
5: geneeqp2002s:*
6: geneeqp2003as:*
7: geneeqp2003bs:*
8: geneeqp2004s:*
9: geneeqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3030	99.1	639	2	AA117089 Bacillus
2	3022	98.8	640	2	AA117090 Bacillus
3	3022	98.8	640	8	ADM40773 Alkaline
4	3022	98.8	640	8	ADSI14437 Bacillus
5	3022	98.8	641	8	ADT49643 Bacillus
6	3021	98.8	640	2	AA117091 Bacillus
7	3017	98.6	639	2	AA117087 An alkali
8	3007	98.3	640	2	AA117088 An alkali
9	2723	89.0	641	2	AAW89547 Bacillus
10	2408	78.7	636	2	AAW89548 Bacillus
11	2155	70.4	434	5	AAW50080 Bacillus
12	2155	70.4	434	5	AAW50081 Bacillus
13	2155	70.4	434	7	ADY33778 Bacillus
14	2155	70.4	434	7	ADZ51758 Mutant Ba
15	2155	70.4	434	7	ADZ51757 Wild-type
16	2155	70.4	434	8	ADL25802 Bacillus
17	2155	70.4	434	8	ADM40779 Alkaline
18	2155	70.4	434	8	ADM40771 Mature al
19	2155	70.4	434	8	ADM40780 Alkaline
20	2155	70.4	434	8	ADM41427 Bacillus
21	2150	70.3	434	8	ADT49604 Bacillus
22	2147	70.2	434	8	ADSI14441 Bacillus
23	2147	70.2	434	8	ADSI14438 Bacillus
24	2147	70.2	434	8	ADSI14442 Bacillus

25	2146	70.2	434	8	ADSI14439 Bacillus
26	2140	70.0	434	8	ADSI14440 Bacillus
27	2138	69.9	434	8	ADSI14443 Bacillus
28	2133	69.7	434	8	ADSI14444 Bacillus
29	2118	69.2	434	8	ADSI14445 Alkaline
30	2103	68.7	436	8	ADM40787 Bacillus
31	2082	68.1	434	5	AAW50085 Bacillus
32	2082	68.1	434	7	ADZ51762 Mutant Ba
33	2064.5	67.5	433	8	ADM40784 Alkaline
34	2064.5	67.5	433	8	ADSI2082 Bacillus
35	2060.5	67.4	433	5	AAW50083 Bacillus
36	2060.5	67.4	433	5	AAW50086 Bacillus
37	2060.5	67.4	433	7	ADZ51763 Mutant Ba
38	2060.5	67.4	433	8	ADM40785 Alkaline
39	2060.5	67.4	433	8	ADSI2010 Bacillus
40	2060.5	67.4	433	8	ADSI2078 Bacillus
41	2060.5	67.4	433	8	ADSI2016 Bacillus
42	2060.5	67.4	433	8	ADSI2011 Bacillus
43	2060.5	67.4	433	8	ADSI2001 Bacillus
44	2059.5	67.3	433	8	ADSI2062 Bacillus
45	2058.5	67.3	433	8	ADSI2054 Bacillus

ALIGNMENTS

RESULT 1	AA117089	standard; protein; 639 AA.
ID	AA117089	standard; protein; 639 AA.
XX	AA117089;	
AC	AA117089;	
XX	20-MAR-2003 (revised)	
DT	21-JUL-1999 (first entry)	
XX	Bacillus alkaline protease.	
DE	Bacillus alkaline protease.	
XX	Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;	
KW	washing composition; oxidizing agent.	
XX	Bacillus sp.	
OS	WO9918218-A1.	
XX	15-APR-1999.	
PD	07-OCT-1998; 98WO-JP004528.	
XX	07-OCT-1997; 97JP-00274570.	
PF	(KAOS) KAO CORP.	
XX	Takaiwa M, Okuda M, Sasaki K, Kubota H, Hitomi J, Kageyama Y;	
PI	Shikata S, Nomura M,	
XX	WPI. 1999-287736/27.	
DR	N-PSDB; AA137277.	
XX	Alkali protease from Bacillus used in washing powders.	
PT	Disclosure; Page 53-58; 71pp; Japanese.	
XX	The invention relates to alkaline proteases produced by strains of	
CC	Bacillus. The proteases ability to digest casein is not inhibited by	
CC	oleic acid and they have a high stability to oxidizing agents. The	
CC	alkaline protease of the invention has the following properties: (a) it	
CC	is active over the pH range 4-13 and has at least 80% of its optimum	
CC	activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is	
CC	stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)	
CC	its ability to digest casein is not inhibited by oleic acid; (e) it has	
CC	molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be	
CC	used as enzymes in washing compositions for use in automatic dishwashers	
CC	and for washing clothes. The stability to oxidizing agents allows the	

CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX

Sequence 639 AA;

Query Match 99.1%; Score 3030; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 3.4e-237;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
QY 1 MRKKVFLSVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTITDXXGFSKXQTGAAPL 60
DB 1 MRKKVFLSVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTITDXXGFSKXQTGAAPL 60
QY 61 LVSEENVKLKXGKXKLETPVANNKLIHQFNPIIEETKQKLETKGAKILDYIPYAYI 120
DB 61 LVSEENVKLKXGKXKLETPVANNKLIHQFNPIIEETKQKLETKGAKILDYIPYAYI 120
QY 121 VVEYGDVXSXXXIEHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 180
DB 121 VVEYGDVXSXXXIEHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 180
QY 181 IEKIAQXXXNDVYITTAPEYVNDVARGIVKADVAQSSYGLYGGQIIVAVDTGLDT 240
DB 181 IEKIAQXXXNDVYITTAPEYVNDVARGIVKADVAQSSYGLYGGQIIVAVDTGLDT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNANDTNGHGTAVAGSYGNGTNGMAPOANLVFQ 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNANDTNGHGTAVAGSYGNGTNGMAPOANLVFQ 300
QY 301 SJMDSXGIGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTYDTSRVDDYVRKNDT 360
DB 301 SJMDSXGIGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTYDTSRVDDYVRKNDT 360
QY 361 ILPAAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 420
DB 361 ILPAAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTIIISARSSLPDSSFMANHDSKYAVMGSTSMATPIVAGVNAQLREHPFN 480
DB 421 IKPDVMAFGTIIISARSSLPDSSFMANHDSKYAVMGSTSMATPIVAGVNAQLREHPFN 480
QY 481 RGITPEPSLLKAAIIGAADVGLGYPNGQGWGRVTLDKSLNAVAVNSSLSTSQKATY 540
DB 481 RGITPEPSLLKAAIIGAADVGLGYPNGQGWGRVTLDKSLNAVAVNSSLSTSQKATY 540
QY 541 XETATAGKPELKTISWSDAPASTTASVTLVNDLIVITRPNCTYVVGDPFKXPKXANDG 600
DB 541 XETATAGKPELKTISWSDAPASTTASVTLVNDLIVITRPNCTYVVGDPFKXPKXANDG 600
QY 601 RNNVENVEFINKPQSGTYTIEVQAVNVVPGQXFSIAIYN 639
DB 601 RNNVENVEFINKPQSGTYTIEVQAVNVVPGQXFSIAIYN 639
```

RESULT 2
AA17090
ID AA17090 standard; protein; 640 AA.

```
XX AA17090;
XX
XX 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
XX DE Bacillus alkaline protease.
XX
XX KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX washing composition; oxidizing agent.
XX
XX Bacillus sp.
XX
XX MO9918218-A1.
```

PD 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
PF
XX
XX 07-OCT-1997; 97JP-00274570.
PR
XX
XX (KAOS) KAO CORP.
PA
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
XX WPI: 1999-287736/27.
DR N-PSDB; AAK37278.
XX

Alkali protease from Bacillus used in washing powders.
Discloure; Page 58-63; 71pp; Japanese.

The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-11; (b) after 30 minutes at 40 deg C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX

Sequence 640 AA;

Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
QY 2 RKKKVFSLVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTITDXXGFSKXQTGAAPL 61
DB 3 RKKKVFSLVLSAAAILSTVALKNPSAGAXRFDLDFKGIQTITDXXGFSKXQTGAAPL 62
QY 62 VESENVKLKXGKXKLETPVANNKLIHQFNPIIEETKQKLETKGAKILDYIPYAYI 121
DB 63 VESENVKLKXGKXKLETPVANNKLIHQFNPIIEETKQKLETKGAKILDYIPYAYI 122
QY 122 EYEGDVXSXXXIEHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 181
DB 123 EYEGDVXSATSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAXALDTKXNKKEVQLRG 182
QY 182 EXIAQXXXNDVYITTAPEYVNDVARGIVKADVAQSSYGLYGGQIIVAVDTGLDTG 241
DB 183 EXIAQFASISNDVYITTAPEYVNDVARGIVKADVAQSSYGLYGGQIIVAVDTGLDTG 242
QY 242 RNDSSMHEAFRGKITLALYALGRTNANDTNGHGTAVAGSYGNGTNGMAPOANLVFOS 301
DB 243 RNDSSMHEAFRGKITLALYALGRTNANDTNGHGTAVAGSYGNGTNGMAPOANLVFOS 302
QY 302 IMDSXGIGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTYDTSRVDDYVRKNDT 361
DB 303 IMDSXGIGLPSNLQTLFSGQXASAGARIHTNSWGAANVGAVTYDTSRVDDYVRKNDT 362
QY 362 LPAAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 421
DB 363 LPAAGNEKPNNGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR 422
QY 422 KPDVMAFGTIIISARSSLPDSSFMANHDSKYAVMGSTSMATPIVAGVNAQLREHPFN 481
DB 423 KPDVMAFGTIIISARSSLPDSSFMANHDSKYAVMGSTSMATPIVAGVNAQLREHPFN 482
QY 482 GITPEPSLLKAAIIGAADVGLGYPNGQGWGRVTLDKSLNAVAVNSSLSTSQKATYX 541
```

Db 483 GTPKPSLLKALLIAGADIGLGYPNGNGMGRVTLDKSLNVAYNNESSLSSTOKATYS 542
QY 542 PTAATGKPLKISLWSDAPASTTASTVTLVNDLDTLTAPNGTXYVGNDRPKXPKXNDGR 601
Db 543 PTAATGKPLKISLWSDAPASTTASTVTLVNDLDTLTAPNGTXYVGNDRPKXPKXNDGR 602
QY 602 NNVENVFINKPQSGTYTIEVQAYNVPVGPQXPSLAIVN 639
Db 603 NNVENVFINAPOSCTTIEVQAYNVPVGPQXPSLAIVN 640

RESULT 3
ADM40773
ID ADM40773 standard; protein; 640 AA.
XX
AC ADM40773;
XX
DT 01-JUL-2004 (first entry)
XX
DE Alkaline protease from Bacillus sp. KSM-KP43.
XX
KM alkaline protease; laundry detergent; bleaching agent; detergent;
XX denture-cleaning agent; enzyme.
XX
OS Bacillus sp.; KSM-KP43.
XX
PN US2004072321-A1.
XX
PD 15-APR-2004.
XX
PF 09-JUN-2003; 2003US-00456479.
XX
PR 26-JUN-2002; 2002JP-00186387.
XX 18-OCT-2002; 2002JP-00304232.
XX
XX (KAOS) KAO CORP.
XX
PI Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;
XX
DR WPI, 2004-328572/30.
XX N-PSDB; ADM40772.
XX
PT New alkaline protease having a mutant prepro sequence where amino acid
PT residues at positions 52, 75 and 142 are substituted with another amino
PT acid residue, useful as enzyme component of laundry detergents, or
PT bleaching agents.
XX
XX Discloure; SEQ ID NO 4; 29pp; English.
XX
XX The invention relates to an alkaline protease having a prepro sequence.
XX The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
XX acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
XX acid residues at: (a) position 52 is substituted by aspartic acid or
XX arginine; (b) position 75 is substituted by alanine or arginine; and (c)
XX position 142 is substituted by lysine. The alkaline protease is useful as
XX an enzyme which can be incorporated into laundry detergents, bleaching
XX agents, detergents for cleaning hard surfaces or drainpipes, denture-
XX cleaning agents, and detergents for sterilizing medical apparatus. The
XX present sequence represents alkaline protease from Bacillus sp. KSM-KP43.
XX
SQ Sequence 640 AA;

Query Match 98.8%; Score 3022; DB 8; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKGFVLSVLSAAATLSTVALKNPSAGXARXFDLPKGIQTDTDXGSPKXOTGAATL 61
Db 3 KKKGVFLSVLSAAATLSTVALSNPSAGXARNFDLPKGIQTDTDAKGFSPKXOTGAATL 62
QY 62 VESENVKLLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 121
Db 63 VESENVKLLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 122

QY 122 EYEGDVKXXXXXIEHVESVPEPLPKXIDPOLPTKGASLYVAKXALDTKXNKEVOLRGI 181
Db 123 EYEGDVKASTSTTEHVESVPEPLPYRIDPOLPTKGASLYVAKXALDTKXNKEVOLRGI 182
QY 182 EXIAQKXNDVXYITAKPEYKVMNDVARGIKADVAQSSYGLYGQGIIVAAVADTGLDYG 241
Db 183 EQIAQPAISNDVLYITAKPEYKVMNDVARGIKADVAQSSYGLYGQGIIVAAVADTGLDYG 242
QY 242 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTIVAGSVLNGXNTKGMAPQANLVFQS 301
Db 243 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTIVAGSVLNGXNTKGMAPQANLVFQS 302
QY 302 IMDSXGGLGSLPSNIQTLFSAQXKAGARHTMSKGAANVAGATTDSRNDVYRKNDMTI 361
Db 303 IMDSGGLGSLPSNIQTLFSAQXKAGARHTMSKGAANVAGATTDSRNDVYRKNDMTI 362
QY 362 LPAAGNEXPGGTISAPGTAKNAITVGATENLRPSRGSYADNINHYAOPSSRGPTDGR 421
Db 363 LPAAGNEXPGGTISAPGTAKNAITVGATENLRPSRGSYADNINHYAOPSSRGPTDGR 422
QY 422 KPDVAPGTXYLSASSSLAPDSSFWANHDSKYAYMGTSMATPIVAGNVAQLRHFVKOR 481
Db 423 KPDVAPGTXYLSASSSLAPDSSFWANHDSKYAYMGTSMATPIVAGNVAQLRHFVKOR 482
QY 482 GTPKPSLLKALLIAGADXGLGYPNGNGMGRVTLDKSLNVAYNNESSLSSTOKATYS 541
Db 483 GTPKPSLLKALLIAGADIGLGYPNGNGMGRVTLDKSLNVAYNNESSLSSTOKATYS 542
QY 542 PTAATGKPLKISLWSDAPASTTASTVTLVNDLDTLTAPNGTXYVGNDRPKXPKXNDGR 601
Db 543 PTAATGKPLKISLWSDAPASTTASTVTLVNDLDTLTAPNGTXYVGNDRPKXPKXNDGR 602
QY 602 NNVENVFINKPQSGTYTIEVQAYNVPVGPQXPSLAIVN 639
Db 603 NNVENVFINAPOSCTTIEVQAYNVPVGPQXPSLAIVN 640

RESULT 4
ADS14437
ID ADS14437 standard; protein; 640 AA.
XX
XX ADS14437;
XX
AC ADS14437;
XX
DT 30-DEC-2004 (first entry)
XX
XX Bacillus alkaline protease KP43.
XX
XX protease; enzyme; alkaline protease; laundry detergent; KP43.
XX
XX Bacillus sp.; KSM-KP43.
XX
XX Key Location/Qualifiers
XX Peptide 1..206
XX Protein 207..640
XX
XX BP1466962-A1.
XX
XX 13-OCT-2004.
XX
XX 08-APR-2004; 2004BP-00008604.
XX
XX 10-APR-2003; 2003JP-00106709.
XX
XX (KAOS) KAO CORP.
XX
XX Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;
XX
XX WPI, 2004-711313/70.
XX
XX N-PSDB; ADS14428.
XX
XX New engineered alkaline protease, useful particularly in laundry
XX detergents, comprising specified amino acids at particular positions.
XX
XX

XX Disclosure; SEQ ID NO 1; 31pp; English.
PS
XX
CC The invention relates to a novel alkaline protease. The new alkaline
CC protease comprises an amino acid sequence in which one or more amino acid
CC residues selected from those located at 7 specific positions within
CC ADS14427, or at positions corresponding to these positions are: position
CC 15 (histidine), position 16 (threonine or glutamine), position 166
CC (glycine), position 167 (valine), position 187 (serine), position 346
CC (arginine), and position 405 (aspartic acid). The alkaline protease is
CC useful in industry particularly in laundry detergents, but also e.g. in
CC fibre modifying agents, leather processing agents, cosmetic compositions,
CC bath additives, food-modifying agents, and pharmaceuticals. The present
CC sequence represents the wild-type Bacillus sp. KSM-KP43 alkaline
CC protease, KP43. The sequence is shown in the sequence listing as part of
CC SEQ ID NO:1
XX
SQ Sequence 640 AA;
Query Match 98.8%; Score 3022; DB 8; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 2 RKKKVFSLVSAAILSTVALNPSAGKARXFDLDFKGIQTTTDXKGFSGKQGTGAAPL 61
DB 3 KKKKVFSLVSAAILSTVALSNPSAGARNFDLDFKGIQTTTDAKGFSGKQGTGAAPL 62
QY 62 VSEBNVKKLKKGLKKLETPANNKLIHXQPNGBILEETKQKLEXTGAKLIDYIPVAYIV 121
DB 63 VSEBNVKKLPKGLQKLETPANNKLIHQPNGBILEETKQKLEXTGAKLIDYIPVAYIV 122
QY 122 EYEGDVYKXXXXLIEHVESVEPYLPYXXIDPOLFTKGSXLYVAXXLDPTKQXKEVOLGI 181
DB 123 EYEGDVKASNTSTIEHVESVEPYLPYRIPQLFTKGSSELVQAVALDTQKKEVOLGI 182
QY 182 EKIAQXXXXNDVYITTAKEPYKMDVARGIVKADVAQSSYGLYGGQIVAAVADTGLDTG 241
DB 183 EOIAPALISNDVLYITAKEPYKMDVARGIVKADVAQSSYGLYGGQIVAAVADTGLDTG 242
QY 242 RNDSMHEAFRGKITLALYALGRTNNANDNGHGTAVAGSVLGNKTNKGMAPQANLVFQS 301
DB 243 RNDSMHEAFRGKITLALYALGRTNNANDNGHGTAVAGSVLGNKTNKGMAPQANLVFQS 302
QY 302 IMDSXGGLGGLPSNLQTLFSGQXKSGARLHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361
DB 303 IMDSGGGLGGLPSNLQTLFSGQAVSAGARLHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 362
QY 362 LPAAGNHEKPNNGGITISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 421
DB 363 LPAAGNHEKPNNGGITISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 422
QY 422 KPDVWAPGTILISARSLAPDSSFWANHDSKYAVNGGTSMATPTIYAGVAVOLREHFVNR 481
DB 423 KPDVWAPGTILISARSLAPDSSFWANHDSKYAVNGGTSMATPTIYAGVAVOLREHFVNR 482
QY 482 GITPSPSLIKKALIGAADIGLGYNGNGQWGRVTLDSLNVAAVYNESSSLSTOKATYS 541
DB 483 GITPSPSLIKKALIGAADIGLGYNGNGQWGRVTLDSLNVAAVYNESSSLSTOKATYS 542
QY 542 FTATAGKELKLSLVWSDAPASTTASVTLVNDLVTTPNGTXYVNDGFXKPKXKNDGR 601
DB 543 FTATAGKELKLSLVWSDAPASTTASVTLVNDLVTTPNGTXYVNDGFXKPKXKNDGR 602
QY 602 NNVENVFINKFOSGTYTIEVOAYNVFVGPOKFSLAIVN 639
DB 603 NNVENVFINKFOSGTYTIEVOAYNVFVGPOKFSLAIVN 640
RESULT 5
ADT49613
XX ADT49613 standard; protein; 641 AA.
XX
AC ADT49613;

XX 30-DEC-2004 (first entry)
DT
XX Bacillus alkaline protease.
XX
DB Alkaline protease; detergent; fiber modification; leather processing;
KW cosmetic; bath additives; food-modification; pharmaceutical; enzyme.
XX Bacillus sp. KSM-KP43.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..206
FT /note= "signal peptide"
FT Protein 207..641
FT /note= "specifically claimed mature protein (SEQ ID 1)"
XX
XX EP1466970-A1.
XX
XX PD 13-OCT-2004.
XX
XX PF 08-APR-2004; 2004EP-00008605.
XX
XX PR 10-APR-2003; 2003JP-00106708.
XX
XX PA (KAOS) KAO CORP.
XX
XX Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T,
XX
XX WPI; 2004-711317/70.
XX
XX DR N-PSDB; ADT49605.
XX
XX PT New engineered alkaline protease with improved activity and thermal
XX stability, useful particularly in detergents such as laundry detergents.
XX
XX Disclosure; Page 19-25; 35pp; English.
XX
XX The invention relates to an alkaline protease and its encoding gene. The
XX alkaline protease can be expressed by standard recombinant methodology.
XX The alkaline protease is useful in the industry particularly in
XX detergents such as laundry detergents, but also in fiber modifying
XX agents, leather processing agents, cosmetic compositions, bath additives,
XX food-modifying agents, and pharmaceuticals. The enzyme has good activity
XX and thermal stability. The present sequence represents an alkaline
XX protease from Bacillus sp. KSM-KP43.
XX
XX Sequence 641 AA;
SQ
Query Match 98.8%; Score 3022; DB 8; Length 641;
Best Local Similarity 93.3%; Pred. No. 1.5e-236;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
QY 2 RKKKVFSLVSAAILSTVALNPSAGKARXFDLDFKGIQTTTDXKGFSGKQGTGAAPL 61
DB 3 KKKKVFSLVSAAILSTVALSNPSAGARNFDLDFKGIQTTTDAKGFSGKQGTGAAPL 62
QY 62 VSEBNVKKLKKGLKKLETPANNKLIHXQPNGBILEETKQKLEXTGAKLIDYIPVAYIV 121
DB 63 VSEBNVKKLPKGLQKLETPANNKLIHQPNGBILEETKQKLEXTGAKLIDYIPVAYIV 122
QY 122 EYEGDVYKXXXXLIEHVESVEPYLPYXXIDPOLFTKGSXLYVAXXLDPTKQXKEVOLGI 181
DB 123 EYEGDVKASNTSTIEHVESVEPYLPYRIPQLFTKGSSELVQAVALDTQKKEVOLGI 182
QY 182 EKIAQXXXXNDVYITTAKEPYKMDVARGIVKADVAQSSYGLYGGQIVAAVADTGLDTG 241
DB 183 EOIAPALISNDVLYITAKEPYKMDVARGIVKADVAQSSYGLYGGQIVAAVADTGLDTG 242
QY 242 RNDSMHEAFRGKITLALYALGRTNNANDNGHGTAVAGSVLGNKTNKGMAPQANLVFQS 301
DB 243 RNDSMHEAFRGKITLALYALGRTNNANDNGHGTAVAGSVLGNKTNKGMAPQANLVFQS 302
QY 302 IMDSXGGLGGLPSNLQTLFSGQXKSGARLHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361

```

Db      303  IMDSGGGLGGLPSNLQTLFSGQAYSAGARIHTNSWGAAVNGAYTTDSRNVDDYVRKNDMTI 362
Qy      362  LPAAGNEXRNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHYAOFSSRGPTDGR1 421
      363  LPAAGNEXRNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHYAOFSSRGPTDGR1 422
Qy      422  KPDVMAPGTXIIISARSSLPDSSPFMANHDSKTYAVMGTSMAITIVAGNVQOLBEHFVKR 481
      423  KPDVMAPGTXIIISARSSLPDSSPFMANHDSKTYAVMGTSMAITIVAGNVQOLBEHFVKR 482
Qy      482  GITPKPSLLKALILAGADXLGYPNGNQGMRVTLDKSLNVAAYNESSXLSTSOQATYX 541
      483  GITPKPSLLKALILAGADIGLGPNGNQGMRVTLDKSLNVAAYNESSXLSTSOQATYX 542
Db      542  FTATAGKPLKISLVMSDAPASTTASVTLVNDLDLVTAPNGTXYYGNDFFXXPKXNWDGR 601
      543  FTATAGKPLKISLVMSDAPASTTASVTLVNDLDLVTAPNGTXYYGNDFFXXPKXNWDGR 602
Qy      602  NNVENVFIXXPOSGTITIEVOAYNVPGPOXPSLAIYN 639
      603  NNVENVFIXXPOSGTITIEVOAYNVPGPOXPSLAIYN 640
Db

RESULT 6
AA17091
ID  AA17091 standard; protein; 640 AA.
XX
AC  AA17091;
XX
DT  20-MAR-2003 (revised)
DT  21-JUL-1999 (first entry)
XX
DE  Bacillus alkaline protease.
XX
KW  Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX  washing composition; oxidising agent.
XX
OS  Bacillus sp.
XX
PN  W09918218-A1.
XX
PD  15-APR-1999.
XX
PF  07-OCT-1998; 98WO-JP004528.
XX
PR  07-OCT-1997; 97JP-00274570.
XX
PA  (KAOS ) KAO CORP.
XX
PI  Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI  Shikata S, Nomura M;
XX
DR  MPI; 1999-287736/27.
XX  N-PSDB; AAX37279.
PT  Alkali protease from Bacillus used in washing powders.
PS  Disclosure; Page 63-68; 71pp; Japanese.
XX
CC  The invention relates to alkaline proteases produced by strains of
CC  oleic acid and they have a high stability to oxidising agents. The
CC  alkaline protease of the invention has the following properties: (a) it
CC  is active over the pH range 4-13 and has at least 80% of its optimum
CC  activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC  stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC  its ability to digest casein is not inhibited by oleic acid; (e) it has
CC  molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC  used as enzymes in washing compositions for use in automatic dishwashers
CC  and for washing clothes. The stability to oxidising agents allows the
CC  enzyme to be an effective component of washing compositions including
CC  bleaches. The present sequence represents an alkaline protease. (Updated
CC  on 20-MAR-2003 to correct DR field.)

```

```

XX      XX      Sequence 640 AA;
SQ      Query Match          98.8%; Score 3021; DB 2; Length 640;
      Best local Similarity 93.3%; Pred. No. 1, 8e-236;
      Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

Qy      2  RKKRVFLSTLSAAIISTVALKNPSAGKARXPDLPKGIQTTHDXGFSKQKGTGAAPL 61
      3  KKKRVFLSTLSAAIISTVALKNPSAGKARXPDLPKGIQTTHDXGFSKQKGTGAAPL 62
Db      62  VSEENVFKLKKGKKLLETVPANNKLIHQFNGPILESTKQJEXTGAKLIDYVDAAYIT 121
      63  VSEENVFKLKKGKKLLETVPANNKLIHQFNGPILESTKQJEXTGAKLIDYVDAAYIT 122
Qy      122  EYEGDVXKXXXXIIEHVESVEPYLPYXXIDIPOLFTKASLYLVKXALDTRQXNKEVOLRG1 181
      123  EYEGDVXKATSTIEHVESVEPYLPYRIDPOLFTKASLYLVKXALDTRQXNKEVOLRG1 182
Db      182  EXIAQKKXNDYXYITAKPEYKVMNDVANGIYKADVAQSSYGLYGGQGIYAVADTGLDTG 241
      183  EQIAQPAISNDVLYITAKPEYKVMNDVANGIYKADVAQSSYGLYGGQGIYAVADTGLDTG 242
Qy      242  RNDSSMHEAPRGKITLALYALGRTNNANDNGHGHVAGSVLNGXTNKGMAPOANTVPQS 301
      243  RNDSSMHEAPRGKITLALYALGRTNNANDNGHGHVAGSVLNGXTNKGMAPOANTVPQS 302
Qy      302  IMDSXGGLGGLPSNLQTLFSGQAYSAGARIHTNSWGAAVNGAYTTDSRNVDDYVRKNDMTI 361
      303  IMDSGGGLGGLPSNLQTLFSGQAYSAGARIHTNSWGAAVNGAYTTDSRNVDDYVRKNDMTI 362
Db      362  LPAAGNEXRNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHYAOFSSRGPTDGR1 421
      363  LPAAGNEXRNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHYAOFSSRGPTDGR1 422
Qy      422  KPDVMAPGTXIIISARSSLPDSSPFMANHDSKTYAVMGTSMAITIVAGNVQOLBEHFVKR 481
      423  KPDVMAPGTXIIISARSSLPDSSPFMANHDSKTYAVMGTSMAITIVAGNVQOLBEHFVKR 482
Qy      482  GITPKPSLLKALILAGADXLGYPNGNQGMRVTLDKSLNVAAYNESSXLSTSOQATYX 541
      483  GITPKPSLLKALILAGADIGLGPNGNQGMRVTLDKSLNVAAYNESSXLSTSOQATYX 542
Db      542  FTATAGKPLKISLVMSDAPASTTASVTLVNDLDLVTAPNGTXYYGNDFFXXPKXNWDGR 601
      543  FTATAGKPLKISLVMSDAPASTTASVTLVNDLDLVTAPNGTXYYGNDFFXXPKXNWDGR 602
Qy      602  NNVENVFIXXPOSGTITIEVOAYNVPGPOXPSLAIYN 639
      603  NNVENVFIXXPOSGTITIEVOAYNVPGPOXPSLAIYN 640
Db

RESULT 7
AA17087
ID  AA17087 standard; protein; 639 AA.
XX
AC  AA17087;
XX
DT  20-MAR-2003 (revised)
DT  21-JUL-1999 (first entry)
XX
DE  An alkaline protease sequence from Bacillus species.
XX
KW  Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX  washing composition; oxidising agent.
XX
OS  Bacillus sp.
XX
FH  Key      Location/Qualifiers
FH  Misc-difference 1..639
FH  FT        /note="all residues indicated as Xaa are arbitrary amino
FT        acids"

```

```

PN MO918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX Shikata S, Nomura M;
XX
XX WPI; 1999-287736/27.
XX
XX N-PSDB; AAX37278.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Claim 3; Page 47-50; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
XX Bacillus. The proteases ability to digest casein is not inhibited by
XX oleic acid and they have a high stability to oxidising agents. The
XX alkaline protease of the invention has the following properties: (a) it
XX is active over the pH range 4-13 and has at least 80% of its optimum
XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg C it is
XX stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
XX its ability to digest casein is not inhibited by oleic acid; (e) it has
XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX used as enzymes in washing compositions for use in automatic dishwashers
XX and for washing clothes. The stability to oxidising agents allows the
XX enzyme to be an effective component of washing compositions including
XX bleaches. The present sequence represents an alkaline protease of the
XX invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
XX Sequence 639 AA;
SQ
Query Match 98.6%; Score 3017; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.8e-236;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRKKVPLSLVLSAAAILSTVALXNPSAGAXXFDLDFKGIQTTTDXGFSKQXQTGAAP 60
DB 1 MRKKVPLSLVLSAAAILSTVALXNPSAGAXXFDLDFKGIQTTTDXGFSKQXQTGAAP 60
QY 61 LVESNVKLXKGLKXKLETPVANNKLTHTXQFNGPILLETQKXLETGAKIIDYIPDAYI 120
DB 61 LVESNVKLXKGLKXKLETPVANNKLTHTXQFNGPILLETQKXLETGAKIIDYIPDAYI 120
QY 121 VEYEGDVASXXXLIEHVESVEPYLPXYXIDLPQFTKGSXILVKAXALDTKOKNKEVOLRG 180
DB 121 VEYEGDVASXXXLIEHVESVEPYLPXYXIDLPQFTKGSXILVKAXALDTKOKNKEVOLRG 180
QY 121 VEYEGDVASXXXLIEHVESVEPYLPXYXIDLPQFTKGSXILVKAXALDTKOKNKEVOLRG 180
DB 121 VEYEGDVASXXXLIEHVESVEPYLPXYXIDLPQFTKGSXILVKAXALDTKOKNKEVOLRG 180
QY 181 IEXIAQXXXSNDVXTTAKPEYKVNNDVARGIVKADVAQSSYGLGQGIIVAVADTGLDT 240
DB 181 IEXIAQXXXSNDVXTTAKPEYKVNNDVARGIVKADVAQSSYGLGQGIIVAVADTGLDT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNNDNTGHHVAGSVLGNKXTTKGAAPQANLVFQ 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNNDNTGHHVAGSVLGNKXTTKGAAPQANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSGQXGAGARIHTNSWGAIVNGAYTTDSRNDVDYAKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSGQXGAGARIHTNSWGAIVNGAYTTDSRNDVDYAKNDMT 360
QY 361 ILFPAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFSYADNINHTVAQFSSRGPYKGR 420
DB 361 ILFPAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFSYADNINHTVAQFSSRGPYKGR 420
QY 421 IKPDVMAFGTIIISARSSLAPDSSFWANHDSKYAMGTSMATPIVAGNVQLEBHFVKN 480
DB 421 IKPDVMAFGTIIISARSSLAPDSSFWANHDSKYAMGTSMATPIVAGNVQLEBHFVKN 480

```

```

QY 481 RGITPKPSLLKRALIAGAADGICYPNGNQGWRVTLDKSLANVAVNNESSXILSTSQKATY 540
DB 481 RGITPKPSLLKRALIAGAADGICYPNGNQGWRVTLDKSLANVAVNNESSXILSTSQKATY 540
QY 541 XFTATAGKPLKISLWSDAPASTTASVTLVNDLVLVITAPNGTYVNGNDFXPPXXMDG 600
DB 541 XFTATAGKPLKISLWSDAPASTTASVTLVNDLVLVITAPNGTYVNGNDFXPPXXMDG 600
QY 601 RNNVENFVFNKPSGTYTIEVOAVNVPGVPOXFSLSAIYN 639
DB 601 RNNVENFVFNKPSGTYTIEVOAVNVPGVPOXFSLSAIYN 639
XX
XX RESULT 8
XX AAY17088
XX ID AAY17088 standard; protein; 640 AA.
XX
XX AAY17088;
XX
XX 20-MAR-2003 (revised)
XX 21-JUN-1999 (first entry)
XX
XX An alkaline protease sequence from Bacillus species.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX washing composition; oxidising agent.
XX
XX Bacillus sp.
XX
XX Key Location/Qualifiers
XX Misc-difference 1..640
XX FT /note="all residues indicated as Xaa are arbitrary amino
XX FT acids"
XX
XX MO918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX Shikata S, Nomura M;
XX
XX WPI; 1999-287736/27.
XX
XX N-PSDB; AAX37278.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Claim 3; Page 50-53; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
XX Bacillus. The proteases ability to digest casein is not inhibited by
XX oleic acid and they have a high stability to oxidising agents. The
XX alkaline protease of the invention has the following properties: (a) it
XX is active over the pH range 4-13 and has at least 80% of its optimum
XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg C it is
XX stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
XX its ability to digest casein is not inhibited by oleic acid; (e) it has
XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX used as enzymes in washing compositions for use in automatic dishwashers
XX and for washing clothes. The stability to oxidising agents allows the
XX enzyme to be an effective component of washing compositions including
XX bleaches. The present sequence represents an alkaline protease of the
XX invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
XX Sequence 640 AA;
SQ
Query Match 98.3%; Score 3007; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.5e-235;

```

	Pf	XX	09-JUN-1998;	98WO-US012005.
	Fx	XX	12-JUN-1997;	9TUS-00873479.
	Rr	XX	(NOVO) NOVO NORDISK BIOTECH INC.	
	Pt	XX	Siloma A, Christianson L;	
	Dd	XX	MPI; 1999-080908/07.	
	Xx	XX	N-Psdb; AA682382.	
	Pt	PT	Novel protease from Bacillus subtilis LC20 - useful in laundry and dishwashing detergents and for leather processing.	
	Ps	PS	Claim 7; Page 53-54; 77pp; English.	
	Cc	CC	This is the amino acid sequence of a novel protease of Bacillus sp. JH177 (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene (see AA682382). The entire protein, including the signal peptide and C-terminal region, has 77% identity to alkaline protease Y (see AA69548) from Bacillus. The invention provides vectors, recombinant host cells and methods for the recombinant production of the protease. The protease is used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for deproteinizing and enhancing the degree of hydrolysis of protein hydrolysates, for flavour development through hydrolysis of proteins, degradation of undesired peptides and in enzymatic synthesis of peptides. It has enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the peroxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins.	
	Cc	CC	Sequence 641 AA:	
	Sq	XX		
			Query Match	89.0%; Score 2723; DB 2; Length 641;
			Best Local Similarity	82.9%; Pred. No. 36-212;
			Matches 532; Conservative	40; Mismatches 66; Indels 4; Gaps 2;
Oy			1 MRK---KKVPLSTLSAAALISTVALKNBPAGAKRFDLPFKIGITOTTHXGFSKQQTGA	57
Dd			1 MKRKSGSRPFLSVLTVALLSSVALSSPETTGANNFBLDFGIETLTLEKATKGKTGK	60
Oy			58 AAPLVSEENVKLKGGKLLKVTPANKKHIFQNFGPILETBKOKLEXTKAKIIDYIPDY	11
Dd			61 ASPLVSSENVKIKPKSIQCKGLEVPVDNKKLIYIQFGSPLIESTQLQLETKAKIIDYIPDY	120
Oy			118 AYIVVEGDVXSXXXXIHVESVEPYLPTYXIDPOLFTKGSASILVKAXALDTCOKNEEQ	177
Dd			121 AYIVEVDGVKAVTNAMLALESVEPELYPLKYKDLPFRGASSELVEVTALDKGRSKSEVR	180
Oy			178 LRGLIEKIQAQQSXNDVXYTTAPERYKVMNDVARIKADVAAQSSTGYGCQIYA VADTG	235
Dd			181 LRGLIEIAQATANDVLVYTPRPXEVEVLNDVARGIKADVANNFGLYGCCQIIA VADTG	240
Oy			238 LDPRGRDSMHEAFREKITAYALARTNNADTNHGHTHVAGSYLGXTTKMKAAPANTL	297
Dd			241 LDTRGRDSSMHEAFREKITAYALARTNNADPNHGHTHVAGSYLVGN-ATNKMAPANTL	299
Oy			298 VFOSINDSKGSLGCLPSNLQTLFSAQKSAGARIFTNSGCAVNAGAYTTDSRYVDYVRKN	355
Dd			300 VFOSINDSGGSLGCLPANIQTILFSAQYSAGARIHTNSGCAVNGAYTTDSRYVDYVRKN	355
Oy			358 DMTILFAAGNEKPNGGTISAPGTACNAITVGATEMLRSPGSYADNIINHWAOFPSSRGPTK	417
Dd			360 DMTILFAAGNEGSGGTISAPGTACNAITVGATEMLRSPGSYADNIINHWAOFPSSRGPTR	419
Oy			418 DGRIKRDWVAPEGYLISAASSILAPOSSFMAHHDSRYAAGTSMATPVVAGNVAQLAEHF	477
Dd			420 DGRIKRDWVAPEGYLISAASSILAPPSSFWAHNHDSRYAAGTSMATPVVAGNVAQLAEHF	479
Oy			478 VKNRGITPPRSLLKALIMGADKDLGFVNGQGKGRVTLDTKSLANVAVNESSXLSTFSQK	537

Db	480	VKNRGCTTPKPSLKKALLIGADVGLGFPNGNCKGRVTLDKSLNAPVNETISPLSTSQK	539
Qy	538	ATYTFATATGKPKLKSILWSDAPASTASVTLVNDLVTAPNGTXYYGNDPEKXPXXN	597
Db	540	ATYSFTAQAGKPKLSILWSDANGSTTASTLTVNDLVTITAPNGTKRYGNDPTAYDNN	599
Qy	598	WNGRRNVENVFINXPOSCTTYTTEVQAINPVPGQKTSIAIVN	639
Db	600	WDGRNNVENVFINAPQSGTYTEVQAINPVFVSPQTSIAIVH	641
RESULT 10			
ID	AAW89548	standard, protein, 636 AA.	
AC	AAW89548;		
XX			
XX	12-APR-1999	(first entry)	
DT			
XX			
DB	Bacillus sp.	alkaline protease Y.	
XX			
KM	Alkaline protease Y; detergent; surfactant; leather processing;		
XX	debittering; flavour.		
OS	Bacillus sp.		
XX			
PN	W09856927-A2.		
PD	17-DEC-1998.		
XX			
XX	09-JUN-1998;	98MO-US012005.	
PE			
XX	12-JUN-1997;	97US-00873479.	
PR			
XX			
PA	(NOVO) NOVO NORDISK BIOTECH INC.		
XX			
PI	Sloma A, Christianson L;		
XX			
XX	WPI; 1999-080908/07.		
DR			
PT	Novel protease from Bacillus subtilis LC20 - useful in laundry and		
PT	dishwashing detergents and for leather processing.		
XX			
XX			
PS	Claim 3; Page 55-56; 77pp; English.		
XX			
CC	This is the amino acid sequence of a Bacillus sp. alkaline protease Y		
CC	that is said to have good alkali and surfactant resistance and improved		
CC	degeneracy. It shows 77% identity to a newly isolated protease (see		
CC	AM89557) of Bacillus sp. JP170 (NCIB 12513). The invention provides		
CC	vector, recombinant host cells and methods for the recombinant		
CC	production of such proteases. The protease are used in laundry and		
CC	dishwashing detergents, for institutional and industrial cleaning, and		
CC	for leather processing, as well as for debittering and enhancing the		
CC	degree of hydrolysis of protein hydrolysates, for flavour development		
CC	through hydrolysis of proteins. Degradation of undesired peptides and in		
CC	enzymatic synthesis of peptides. They have enhanced stability towards		
CC	oxidation under alkaline conditions, e.g. towards bleaching agents of the		
CC	peroxy type. The invention also provides mutant cells in which the		
CC	protease activity is diminished. Such cells can be used for the		
CC	production of heterologous recombinant proteins		
XX			
XX			
SO	Sequence 636 AA;		
Query Match	78.7%;	Score 2408; DB 2; Length 636;	
Best Local Similarity	73.2%;	Pred. No. 1,2e-186;	
Matches	467; Conservative	67; Mismatches 98; Indels 6; Gaps 5;	
Qy	3	KKKVFSLVLSAAITSTVTLNKPSSGAKRXPDLDFKGIQTDDXXGFSKQQTGAAPLV	62
Db	4	KKRVVLSSVSAALIASVWSSPTSQA--DFQVNFNGVK-SLENNASLVKPISSGEASPLV	60
Qy	63	ESENVKLLKGLKKGLQETVPANNKLHIFN-CPILFETKQLETKGAILDIYIPYATIV	121

Db	61	DTEININIPKGIQKCLAEAVQKDNELIVQFPTPGPISEEEKKGLSILGVSILDVYPDYAFIV	120
Qy	122	EYEGDVXSKXXXXIHEHVESVEPYLPXXIDPOLFTGSAKLVAAXLDTQXNKEVQLRG	181
Db	121	QISG-ATKRISTLHSEVENQFPLPKYKIDPELLTKGASGLVQAVIILNTHGENKMFTEGL	179
Qy	182	EXIAQXXKSNNDYXYITAKPEYKVMNDVARGIKVADVAQSSYGLYGGQIWAVALDTLDTG	241
Db	180	DEIVQAAANNDDVLIYSIPKPEYELMDVANGIKVADVAQNNYGLYGGQGLVAVADIGLDIG	239
Qy	242	RUDSSMHEAFRQKITLALVALGRTNNANDTNHGCTHVAGSVLGNKXTNKGMAPOANI.VFQS	301
Db	240	RUDSSMHEAFRQKITLALVALGRTNNASDPNGCHTIVAGSVLGN-ALNKGMAPQANI.VFQS	298
Qy	302	INDSXGGLGGLSNIQTLFSAQXXSAGARIHTMSWGAAYVGAATTDSSRNVDDVRYKDDMTI	361
Db	299	INDSSGGLGGLPSNLNLTLSQAMNAGARLHTMSGAPVNGAATYANSRQVDEYRNNDDTV	358
Qy	362	LEPAENEXPNGGTISAPGTAKNAITVGATENTLRPSFGSYADNIINHVAQFSSRGPTQGR	421
Db	359	LEPAENEGNSTISAPGTAKNAITVGATENTLRPSFGSIADPNHIIAOPSSSGATRDGRI	418
Qy	422	KEDVNAPEGTXYIISARSSLAPOSSFWANHDSKYAYNGGTSMATPIVAGNVAQLREHFVQNR	481
Db	419	KEDVNAPEGTXYIISARSSLAPOSSFWANYSKYAYNGGTSMATPIVAGNVAQLREHFVQNR	478
Qy	482	GITTPPSLILKALINAGADXLGYEPNGQMGWRVTLKDSLVNAVYNESXLSSTQKATYX	541
Db	479	GITTPPSLILKALINAGATDVGLGYSGDQMGWRVTLKDSLVNAVYNENATALATGQKATYX	538
Qy	542	FTATAGKPLKISLVMSDAPASTTASVTLVNDLIDLITAPNGTXYVGNDFKXXKXNDGR	601
Db	539	FOAQGKDELKISLVMTDAPGSTTASVTLVNDLIDLITAPNGKXYVGNDFKSAAYDNNMGR	598
Qy	602	NNVENVFITNXPOSQTYTIEVQAYNPVGPQFSLATVN	639
Db	599	NNVENVFITNAPOSQTYTIEVQAYNPVSGPQFSLATIVH	636
RESULT 11			
AAM50080			
1D	AAM50080	standard; protein; 434 AA.	
XX			
AC	AAM50080;		
XX			
DT	12-AUG-2002	(first entry)	
XX			
DE	Bacillus sp KSM-KP43	alkaline proteasee protein fragment.	
XX			
KW	Alkaline proteasee; detergent; laundry; bleaching; dishwasher.		
XX			
OS	Bacillus sp.		
XX			
FN	EPI209233-A2.		
XX			
PD	29-MAY-2002.		
XX			
PF	22-NOV-2001; 2001EP-00127851.		
XX			
PR	22-NOV-2000; 2000JP-00355166.		
XX			
PR	12-APR-2001; 2001JP-00114048.		
XX			
PA	(KAOS) KAO CORP.		
XX			
PI	Hatada Y, Ogawa A, Kagayama Y, Sato T, Araki H, Sumitomo N,		
XX	Okuda M, Saeki K;		
XX			
DR	WPI; 2002-437518/47.		
XX			
PT	New modified alkaline proteases useful in detergent compositions.		
XX			
PS	Claim 1; Page 10-11; 25pp; English.		

CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency *
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090

XX
 CC
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.5e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVADVAOSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITLALYALGRTN 265
 DB 1 NDVARGIVADVAOSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITLALYALGRTN 60
 QY 266 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFOSIMSGGLGGLPSNLQTLFSGQAYS 325
 DB 61 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFOSIMSGGLGGLPSNLQTLFSGQAYS 120
 QY 326 AGARHTNSWGAAVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 385
 DB 121 AGARHTNSWGAAVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 180
 QY 386 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSGF 445
 DB 181 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSGF 240
 QY 446 WANHDSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADVLGY 505
 DB 241 WANHDSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADVLGY 300
 QY 506 PNGNGMGKRVTLDKSLNVAAYNNESSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 565
 DB 301 PNGNGMGKRVTLDKSLNVAAYNNESSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 360
 QY 566 SYTLVNDLDELVTAPNGTXYVGNDFXKPKXKXMDGRNNVENVFINKPQSGTYTIIYQAYN 625
 DB 361 SYTLVNDLDELVTAPNGTXYVGNDFXKPKXKXMDGRNNVENVFINKPQSGTYTIIYQAYN 420
 QY 626 VEVGPQXPSLAIYN 639
 DB 421 VEVGPQXPSLAIYN 434

RESULT 12
 AAM50081

ID AAM50081 standard; protein; 434 AA.

XX AAM50081;
 XX
 AC
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DB *Bacillus* sp KSM-KP9860 alkaline protease protein fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS *Bacillus* sp.
 XX
 PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001BP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 XX
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.
 XX
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 5; Page 12-13; 25pp; English.

CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency *
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC *Bacillus* sp strain KSM-KP9860 described in the method of the invention

XX
 CC
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.5e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVADVAOSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITLALYALGRTN 265
 DB 1 NDVARGIVADVAOSSYGLYGQGIYAVADTGLDTRNDSSMHEAFRGKITLALYALGRTN 60
 QY 266 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFOSIMSGGLGGLPSNLQTLFSGQAYS 325
 DB 61 NNDTNGHGTHTVAGSVLGNKXTNKGMAPOANLVFOSIMSGGLGGLPSNLQTLFSGQAYS 120
 QY 326 AGARHTNSWGAAVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 385
 DB 121 AGARHTNSWGAAVNGAYTTDSRNVDVYRKNDMTLLFAAGNEXPGGTISAPGTAKNAI 180
 QY 386 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSGF 445
 DB 181 TVGATENTLRPSFGSYADNINHAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSGF 240
 QY 446 WANHDSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADVLGY 505
 DB 241 WANHDSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAGADVLGY 300
 QY 506 PNGNGMGKRVTLDKSLNVAAYNNESSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 565
 DB 301 PNGNGMGKRVTLDKSLNVAAYNNESSLSTSOAKATYXFTATAGKPLKISLWSDAPASTTA 360
 QY 566 SYTLVNDLDELVTAPNGTXYVGNDFXKPKXKXMDGRNNVENVFINKPQSGTYTIIYQAYN 625
 DB 361 SYTLVNDLDELVTAPNGTXYVGNDFXKPKXKXMDGRNNVENVFINKPQSGTYTIIYQAYN 420
 QY 626 VEVGPQXPSLAIYN 639
 DB 421 VEVGPQXPSLAIYN 434

RESULT 13

ID ADY33778 standard; protein; 434 AA.

XX ADY33778;
 XX
 AC
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DB *Bacillus* species alkaline protease.
 XX
 KW mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 XX
 OS *Bacillus* sp.
 XX
 PN EPI347044-A2.
 XX
 PD 24-SEP-2003.
 XX

PF 21-MAR-2003; 2003EP-0006472.
 XX
 XX 22-MAR-2002; 2002JP-00081428.
 PR 06-JUN-2002; 2002JP-00165987.
 PR 18-OCT-2002; 2002JP-00304230.
 PR 18-OCT-2002; 2002JP-00304231.
 XX
 XX (KAOS) KAO CORP.
 XX
 XX Okuda M, Sato[†] H, Saito K, Sumitomo N, Izawa Y, Saeki K;
 PI Kobayashi T, Nomura M;
 DR WPI; 2003-846540/79.
 XX N-PSDB; ADY33779.
 XX
 PT New alkaline protease having specific amino acid residue at a specific
 PT position of its amino acid sequence, useful for producing detergent
 PT compositions, laundry detergent, fiber modifiers, leather-treating agents
 PT or pipe cleaners.
 XX
 PS Claim 1; SEQ ID NO 1; 31pp; English.
 XX
 CC The invention relates to an alkaline protease having a fully defined
 CC sequence of 434 amino acids (I) given in the specification, or an amino
 CC acid sequence at least 80% homology with (I), where an amino acid residue
 CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
 CC selected from 16 amino acid residues. The amino acid residues at the
 CC corresponding positions are selected from: position 65: proline; position
 CC 101: asparagine; position 163: histidine, aspartic acid, phenylalanine,
 CC lysine, asparagine, serine, histidine, leucine, glutamine, threonine or
 CC valine; position 170: valine or leucine; position 171: alanine, glutamic
 CC acid, glycine or threonine; position 273: isoleucine, glycine or
 CC threonine; position 320: phenylalanine, valine, threonine, leucine,
 CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
 CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
 CC acid, arginine or histidine. The alkaline protease is useful for the
 CC production of a detergent composition, such as laundry detergent, fiber
 CC modifiers, leather-treating agents, cosmetic compositions, bath
 CC additives, food modifiers and pharmaceutical compositions. The alkaline
 CC protease may also be used as bleaching detergent, hard surface cleansing
 CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
 CC cleanser for medical tools. The new alkaline protease has a more potent
 CC proteolytic capacity, exhibiting excellent detergency for the removal of
 CC a complex stain, and has high secretion capacity. This sequence
 CC corresponds to the *Bacillus* sp. alkaline protease.
 XX
 XX Sequence 434 AA;
 SQ
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.5e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 206 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTGRDSSMHEAFRGKITALYALGRTN 265
 DB 1 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTGRDSSMHEAFRGKITALYALGRTN 60
 QY 266 NANTNGGTHVAGSVLGGNGXTNKGMAPOANLVPSINDSXGGLGSLNLTQTFEQAXS 325
 DB 61 NANTNGGTHVAGSVLGGNGXTNKGMAPOANLVPSINDSXGGLGSLNLTQTFEQAXS 120
 QY 336 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAAGNEXNGGTSAPGTAKNAI 385
 DB 121 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAAGNEXNGGTSAPGTAKNAI 180
 QY 386 TVGATENTRPFPSGSIADNINRVAQSSRGPTDGRIRKPDVWAPGTXYIIISARSLAPDSSF 445
 DB 181 TVGATENTRPFPSGSIADNINRVAQSSRGPTDGRIRKPDVWAPGTXYIIISARSLAPDSSF 240
 QY 446 WANHSKAYVMGTSMTATPIVAGNVAQRLREHPVKRGTITPKPSLLKALLINGAADXGIGY 505
 DB 241 WANHSKAYVMGTSMTATPIVAGNVAQRLREHPVKRGTITPKPSLLKALLINGAADXGIGY 300
 QY 506 PNGNGMGRVTLTDLKSLNVAAYNBSXSLSTGKATYXFTATAGKPLKLSLVMSDAPASTTA 565

DB 301 PNGNGMGRVTLTDLKSLNVAAYNBSXSLSTGKATYXFTATAGKPLKLSLVMSDAPASTTA 360
 QY 566 SVTLVNDLDTVITAPNGTXYVGNDEPKXPKXKNMDGRNNVENVFITNXPQSGTYTTEVQAYN 625
 DB 361 SVTLVNDLDTVITAPNGTXYVGNDEPKXPKXKNMDGRNNVENVFITNXPQSGTYTTEVQAYN 420
 QY 626 VPVGPQKXSLATVYN 639
 DB 421 VPVGPQKXSLATVYN 434
 RESULT 14
 ADZ51758
 ID ADZ51758 standard; protein, 434 AA.
 XX
 AC ADZ51758;
 XX
 DT 16-JUN-2005 (first entry)
 XX
 DE Mutant *Bacillus* sp. alkali protease, SEQ ID No:2.
 XX
 KW alkali protease; enzyme; surfactant; mutagenesis; muteln.
 XX
 OS *Bacillus* sp.
 OS Synthetic.
 XX
 FN JP2003125783-A.
 XX
 PD 07-MAY-2003.
 XX
 PF 26-OCT-2001; 2001JP-00329472.
 XX
 PR 26-OCT-2001; 2001JP-00329472.
 XX
 PR 26-OCT-2001; 2001JP-00329472.
 XX
 PA (KAOS) KAO CORP.
 XX
 DR WPI; 2003-855669/80.
 XX
 PT New alkali protease useful for preparing detergents comprises
 PT substitution mutations at 251 or 256 position of protease KP43 derived
 PT from *Bacillus* species KSM-KP43.
 XX
 PS Claim 3; SEQ ID NO 2; 16pp; Japanese.
 XX
 CC The invention relates to a mutant alkali protease having an amino acid
 CC deletion at position(s) 251 or 256 in a fully defined sequence given as
 CC SEQ ID No:1 in the specification, or the following amino acid residue
 CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
 CC amino acid residue Lys, Ser Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
 CC Glu at position 256. Also described are: (i) an alkali protease having
 CC the amino acid sequence of SEQ ID No:1, or having 60% or more homology to
 CC SEQ ID No:1, or a deletion of amino acid corresponding to position 251 or
 CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
 CC vector containing the gene, and (iv) a transformed organism containing
 CC the recombinant vector. The alkali protease sequence having 60% or more
 CC homology to SEQ ID No:1 is chosen from a fully defined sequence selected
 CC from SEQ ID Nos 2-7 as given in the specification. The transformed
 CC organism is a microorganism. The mutant alkali protease is useful for
 CC preparing detergents. The mutant alkali protease exhibits high resistance
 CC against oxidizing agent. The mutant alkali protease has high specific
 CC activity. This sequence represents a mutant *Bacillus* sp. alkali protease.
 XX
 XX Sequence 434 AA;
 SQ
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.5e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 206 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTGRDSSMHEAFRGKITALYALGRTN 265
 DB 1 NDVARGIVKADVAAQSSYGLYGQGIIVAVADTGLDTGRDSSMHEAFRGKITALYALGRTN 60

QY 266 NANDTNGHGTAVAGSVTLGNGXTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 325
DB 61 NANDTNGHGTAVAGSVTLGNGXTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 120
QY 326 AGARLHTNSMGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 385
DB 121 AGARLHTNSMGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADNINHVAFSSRGPYKGRIPKDVMAPGTYILSARSSLADSSP 445
DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPYKGRIPKDVMAPGTYILSARSSLADSSP 240
QY 446 WANHSKAYMGSTSMATPIVAGNVAQLREHFKVKNRGITPKPSLLKALIIAGAADGIGY 505
DB 241 WANHSKAYMGSTSMATPIVAGNVAQLREHFKVKNRGITPKPSLLKALIIAGAADGIGY 300
QY 506 PNGNGCMGRVTLDKSLNVAAYVNSSXLSTSQKATYXFTATAGKPLKISLVMSDAPASTTA 565
DB 301 PNGNGCMGRVTLDKSLNVAAYVNSSXLSTSQKATYXFTATAGKPLKISLVMSDAPASTTA 360
QY 566 SYTLVNDLVLVTAPNGTYVGNDFPKXPKXKNWGRNNVENVFINKPQSGTYTIEVOAYN 625
DB 361 SYTLVNDLVLVTAPNGTYVGNDFPKXPKXKNWGRNNVENVFINKPQSGTYTIEVOAYN 420
QY 626 VPVGPOXPSLATVN 639
DB 421 VPVGPOXPSLATVN 434

RESULT 15
AD251757
ID AD251757 standard; protein; 434 AA.
XX
AC AD251757;
XX
DT 16-JUN-2005 (first entry)
XX
DB Wild-type *Bacillus* sp. alkali protease.
XX
OS alkali protease; enzyme; surfactant; mutagenesis.
XX
PN *Bacillus* sp.
XX
PP JP2003125783-A.
XX
PD 07-MAY-2003.
XX
PF 26-OCT-2001; 2P0LJP-00329472.
XX
PR 26-OCT-2001; 200LJP-00329472.
XX
PA (KAOs) KAO CORP.
XX
WI; 2003-856569/80.
XX
PT New alkali protease useful for preparing detergents comprises
PT substitution mutations at 251 or 256 position of protease KP43 derived
PT from *Bacillus* species KSM-KP43.
XX
PS Claim 1, SEQ ID NO 1; 16pp; Japanese.
XX
CC The invention relates to a mutant alkali protease having an amino acid
CC deletion at position(s) 251 or 256 in a fully defined sequence given as
CC SEQ ID No.1 in the specification, or the following amino acid residue
CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
CC amino acid residue Lys, Ser, Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
CC Glu at position 256. Also described are: (i) an alkali protease having
CC the amino acid sequence of SEQ ID No.1, or having 60% or more homology to
CC SEQ ID No.1, or a deletion of amino acid corresponding to position 251 or
CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
CC vector containing the gene, and (iv) a transformed organism containing
CC the recombinant vector. The alkali protease sequence having 60% or more
CC homology to SEQ ID No.1 is chosen from a fully defined sequence selected

CC from SEQ ID Nos 2-7 as given in the specification. The transformed
CC organism is a microorganism. The mutant alkali protease is useful for
CC preparing detergents. The mutant alkali protease exhibits high resistance
CC against oxidizing agent. The mutant alkali protease has high specific
CC activity. This sequence represents wild-type *Bacillus* sp. alkali
CC protease.
SQ Sequence 434 AA;
Query Match 70.4%; Score 2155; DB 7; Length 434;
Best local similarity 96.3%; Pred. No. 2.5e-166;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 206 NDVARGIVADVAOSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGITLVALGRTN 265
DB 1 NDVARGIVADVAOSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGITLVALGRTN 60
QY 266 NANDTNGHGTAVAGSVTLGNGXTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 325
DB 61 NANDTNGHGTAVAGSVTLGNGXTNKGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 120
QY 326 AGARLHTNSMGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 385
DB 121 AGARLHTNSMGAAVNGAYTTDSRNVDYVRKNDMTLLFAAGNEXPNCGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADNINHVAFSSRGPYKGRIPKDVMAPGTYILSARSSLADSSP 445
DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPYKGRIPKDVMAPGTYILSARSSLADSSP 240
QY 446 WANHSKAYMGSTSMATPIVAGNVAQLREHFKVKNRGITPKPSLLKALIIAGAADGIGY 505
DB 241 WANHSKAYMGSTSMATPIVAGNVAQLREHFKVKNRGITPKPSLLKALIIAGAADGIGY 300
QY 506 PNGNGCMGRVTLDKSLNVAAYVNSSXLSTSQKATYXFTATAGKPLKISLVMSDAPASTTA 565
DB 301 PNGNGCMGRVTLDKSLNVAAYVNSSXLSTSQKATYXFTATAGKPLKISLVMSDAPASTTA 360
QY 566 SYTLVNDLVLVTAPNGTYVGNDFPKXPKXKNWGRNNVENVFINKPQSGTYTIEVOAYN 625
DB 361 SYTLVNDLVLVTAPNGTYVGNDFPKXPKXKNWGRNNVENVFINKPQSGTYTIEVOAYN 420
QY 626 VPVGPOXPSLATVN 639
DB 421 VPVGPOXPSLATVN 434

Search completed: April 7, 2006, 17:37:30
Job time : 140.392 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2006, 17:37:51 ; Search time 27.4785 Seconds

(without alignments)
2237.477 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059
Sequence: 1 MKKKVPLSVSAALISLV.....EVOGVNVPYGPQPSLAIYN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518.5	16.9	1743	2 T18279	multidrug resistan
2	477.5	15.6	1905	2 T18267	multidrug resistan
3	350.5	11.5	444	2 B83891	intracellular alka
4	319.5	10.4	442	2 A69587	intracellular alka
5	297.5	9.7	1398	2 T28159	pyrolysate (EC 3.4.
6	286	9.3	806	2 A41341	microbial serine p
7	279.5	9.1	419	1 S25835	subtilisin (EC 3.4
8	276.5	9.0	799	2 G83753	subtilisin-type pr
9	274.5	9.0	420	1 S23407	subtilisin (EC 3.4
10	273.5	8.9	580	2 S11890	serine proteinase
11	268.5	8.8	715	2 JCA908	alkaline serine pr
12	262	8.6	894	2 P68730	cell wall-associat
13	260.5	8.5	1345	2 T29090	surface layer-asso
14	257.5	8.4	378	2 A33973	high-alkaline seri
15	255.5	8.4	381	2 JH0778	subtilisin (EC 3.4
16	255	8.3	513	1 A35742	subtilisin (EC 3.4
17	253.5	8.3	381	1 SUBSI	subtilisin (EC 3.4
18	253	8.3	382	1 SUBSN	subtilisin (EC 3.4
19	251.5	8.2	381	1 SUBSS	subtilisin (EC 3.4
20	250.5	8.2	380	2 A49778	subtilisin (EC 3.4
21	250.5	8.2	601	2 JCA576	serine proteinase
22	250	8.2	382	2 T39780	subtilisin (EC 3.4
23	249.5	8.2	379	1 SUBSCL	subtilisin (EC 3.4
24	249.5	8.2	757	2 C84120	subtilisin-type pr
25	248	8.1	627	2 D75393	serine proteinase,
26	247.5	8.1	381	2 JQ1487	subtilisin (EC 3.4
27	246.5	8.1	402	1 JQ0332	alkaline proteinas
28	242	7.9	1331	2 A72647	probable surface 1
29	239.5	7.8	401	2 I39974	serine proteinase

30	239.5	7.8	534	1 JS0173	alkaline proteinase
31	235	7.7	519	2 S71451	haloalysin R4 (EC 3
32	232	7.6	488	2 A11930	proteinase (import
33	231.5	7.6	1167	1 A35066	streptococcal Csa
34	230	7.5	384	2 JCA802	alkaline proteinase
35	223.5	7.3	409	1 S33905	serine proteinase
36	223	7.3	321	1 S27501	alkaline proteinase
37	220.5	7.2	535	2 B83358	alkaline serine pr
38	219.5	7.2	613	2 S75976	hypothetical prote
39	218.5	7.1	272	2 A23624	subtilisin (EC 3.4
40	218.5	7.1	533	1 JU0146	serine proteinase
41	218.5	7.1	361	2 G83756	subtilisin-type al
42	218	7.1	361	2 G83756	high-alkaline seri
43	216.5	7.1	361	2 A48373	haloalysin (impor
44	216.5	7.1	525	2 G84406	serine proteinase
45	216.5	7.1	645	1 SUBSMP	

ALIGNMENTS

RESULT 1

multidrug resistance transport protein - alime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18279
R:Shaulsky, G.; Loomis, W.F.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z18855
A:Accession: T18279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1743 <SHA>
A:Cross-references: UNIPROT:Q23688; UNIPARC:UP1000013687D; EMBL:U60086; NID:G1399914; PI
A:Gene: tagc

Query Match 16.9%; Score 518.5; DB 2; Length 1743;

Best Local Similarity 24.2%; Pred. No. 2.8e-28;
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY	67	HIXQFNGPLIBETKXKL-----EXTGAKIIDYIPDYAYIV---EYEDVYXK 130
DB	115	YIVQFKRINDTRQRLKPLIGTDIVDEQYQSHIVHYIPDSFLVMTQEGSVLSS 174
QY	131	XXXIHVESVEPY-----LPYXXIDPOLFTKGASXLVK--AXALDT--KQXN 173
DB	175	KEMVSWIGEFERSNKIHLMVNEKSTGLPVYII---LSDSTNSLQKMENTLNSILKSYN 230
QY	174	KEVQLRGI-----EKIAQXXXNDVXYITAKPEYK 203
DB	221	SKVKLTLLNQKKLKSIVCNDESPSPSCSLINSEKLVYQWISQGESN---FIRSEKQ 287
QY	204	VNADVARGIV-----KADVAQSSYGLYGGQGVAAVADPTLDGR--NDS---- 245
DB	288	TANRLSPKVPFGTKDTLVNNDNRVDIP---LRGQQLISTADTGLDSCHPFSQKPI 342
QY	246	---SWHBAFGKITVALGRTNNANDTNGGTHVAGSVLG-----NGXTNKGADPA 295
DB	343	PLNSVNLNHR-KVTVYITTSDDSDKYDGHGTHICGSAAGTPBSSVNISFSLADDA 401
QY	296	NLVFQSIDSKXGLGL--PSNLQTLFQOAXSAGARIHTNSNGA---AVNGAYTTDSRN 349
DB	402	KIAF---FDLASGSSSLTPPSDLKQLYQPLVDAGARVHCDSGVSVEGYSVSOTAS 458
QY	350	VDDVYRKN-DWITLFAENKXPNGSTIS--APGTAKNAITTGATNMLR-----PSGGS 399
DB	459	IDDFLFTHPDFIILAAQN---NEQYLSLTQSTAKNVITVGANQTIHENYLTDDPNYIN 515
QY	400	YADNI----- 404
DB	516	YQSSVDINQELICDPDSRYCMTTAAQCCLESNATTLGLASCCPTLLRKSVIDAANTQPLLY 575

QY 405 --NHVAQSSRGPTKGRIRKPDVMAFGTILSARSLA-----DSSFWAHDSKRYAM 456
DB 576 NENNICSSRSSKGTPTDGRKPKALVAPGEYITARSANGATTQCGGSI--PTNALLA-I 633
QY 457 GGTSMATPIVAGNVAQLREH-----FYKNGITPKPSLLKRAALAGA----- 498
DB 634 SGTSMATSPAAATAATILRLQYLDGYVPTGSIVESNKLOPTGSLLRALMINNAQLNGTFQ 693
QY 499 --ADXLGLGYPNGN-----QGWGRVTLDKSLNVAVNESS----- 530
DB 694 LITSSSIYTPSNQVFNPAQSLVQGWGAIIRMSNMLHVNNNNNSNNNNKTSIDGTRKFDGI 753
QY 531 -----XLISTOKATYXFT-----ATAGKPLK---ISLWMSDAPASTT 564
DB 754 GGLDLRLKPNQWKEBSLSTGQNTSYCTFYKPSSSSSNSGNNIPRVVATLVTWTDPPSYAG 813
QY 565 ASVTLVNDLVI-----TAPNGTYVGNDFPKXXKXWDGRNVEYFVIXP 612
DB 814 AKFNLVNNLDMITMYRDRNGSTIFYSNCGSSFLG-----LAPQDTLNVGIVHNPT 867
QY 613 QSGTITIEVQAVNVPVGPQKFS 634
DB 868 EEMTYRFVAVGATNVPWGPQKFS 889

RESULT 2

T18267
multidrug resistance protein - slime mold (Dictyostelium discoideum)
C1Species: Dictyostelium discoideum
C1Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C1Accession: T18267
R1Shaulsky, G.; Kuspa, A.; Loomis, W.F.
submitted to the EMBL Data Library, January 1995
A1Description: An MDR transporter/serine protease gene is required for prestalk speciali
A1Reference number: Z18850
A1Accession: T18267
A1Status: preliminary; translated from GB/EMBL/DBJ
A1Molecule type: DNA
A1Residues: 1-1905 <SHA>
A1Cross-references: UNIPROT:P54683; UNIPARC:UPI000013687C; EMBL:U20432; NID:g664839; PID
C1Genetics:
A1Gene: tagB

Query Match 15.6%; Score 477.5; DB 2; Length 1905;
Best Local Similarity 24.5%; Pred. No. 2.5e-25;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;
QY 87 HIXQFNGPILBETKQXLE---XTGAKILDYIPDYAVIYEV-----EGDYXS 129
DB 172 YIVQPKDINDBTRQAKBFLGDTITLKBQPFKSHVHYIPIHDSFLVMTKESVLLS 231
QY 130 XXXXIHHVESVBPY-----LPXYI---DPQLFTKASKYL-----V 162
DB 232 SKEMISWIGHEHPSNKHILNYHEKSIQYPVYIILSGTNSLIQWMENTLNSITLTSYNSKV 291
QY 163 KAXALDTKQ-----XNKVQLRGIEIXIAQXXX--NDVXYITAKREYKVM 205
DB 292 KTLTLNQKSLSIYCNDESPSSSSCSLIGSKSIYKWSISEGSESNYIERSEKLDQA 351
QY 206 NDVARGIV-----KADVAGSYGLYGGQIVAVADTGLDTRG---NDS----- 245
DB 352 NRLSPTVIFGTMDKLVNNDRIDIP-----LRGKGQILSLADTGLDGSCHFSDSKYPPIF 406
QY 246 -SMHEAFPGKITLALALGRTNANDTNGHGVAGSVLGNKXTN-----KMAPOANL 297
DB 407 NQVVENHKKVYV---YITYHNDIEDYVNGHGHVCGSAACTPEDSSMAISFSGLATDAKI 463
QY 298 VFGSMDXGGLGIGLPSNLQTLFSQAKSAGARIHTNSMGA-----AVNAGYTTDSNVDY 353
DB 464 APTD-LSSGSSBPPTPEBDSQKPLYDAGARVHDSWGSVSLQGYGYGSDADAGIDAF 522
QY 354 VRK-NDMTILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPFGS-----YADNI- 404

DB 523 LYHPEBSITLAAAGNN--ELFASLAAQATAKNAITVGAEGTAAHVNVSDALEYPSDAN 581
QY 405 -----NHVAQF 410
DB 582 FQRPCLPDKKCYCNTTAKCCSEVSNVKGLQCCPASIKONASDFTTQPPQFENNMGSF 641
QY 411 SSRPTKGRIRKPDVMAFGTILSARSS-----LAPSSFWAHDSKRYAMGTSMATP 464
DB 642 SSKGPTHDGRKPKDIPVAPGEYITARSNGENSTQCGGSI--PNANGLMSISGSMATP 699
QY 465 IVAGNVAQLREH-----YKNGITPKPSLLKRAALI----- 495
DB 700 LATATITLRQYLDGYVPTGSIVESNKLEPTGSLLRALMINNAQLNGTYFWBASSTNP 759
QY 496 AGAADXGLGYPNQGWGRVTLDKSLNVAVNESS-----XLISTOKAT- 539
DB 760 SNAIFEQINGANLIGWGLALRMN---NWLKYSNPPRPPSRWIGIGIGLGRKQKATWKED 816
QY 540 -----YXFT-----ATAGKP-LKISLWMSDAPASTASVTLVNDL- 575
DB 817 SLSSGLNKSICYFTYKPSSSSSSGGGGGETPRIVATLVTWTDPPSYGAKFNLVNNLDDL 876
QY 576 -----VITAPN--GTYYVGNDFPKXXKXWDGRNVEYFVIXPGSTYIEVQAVNVP 627
DB 877 NSDDSIITIGNSGGSIQAPAKVAQ-----DTLANNVEGIIINPTKAMNYKFTIAGTNVP 931
QY 628 VGPQKFS 634
DB 932 IGPQKFS 938

RESULT 3

B83891
intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C
C1Species: Bacillus halodurans
C1Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C1Accession: B83891
R1Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A1Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A1Reference number: A83650; MUID:20512582; PMID:11058132
A1Accession: B83891
A1Status: preliminary
A1Molecule type: DNA
A1Residues: 1-444 <STO>
A1Cross-references: UNIPROT:Q9KBU7; UNIPARC:UPI00000C3D43; GB:AP001513; GB:BA000004; NID
A1Experimental source: strain C-125
C1Genetics:
A1Gene: aprX

Query Match 11.5%; Score 350.5; DB 2; Length 444;
Best Local Similarity 30.2%; Pred. No. 3.2e-17;
Matches 114; Conservative 54; Mismatches 127; Indels 83; Gaps 17;
QY 182 EXIAQXXXNDVXYITAKREYKVMNDVARGIVRA-DVAGSYGLYGGQIVAVADTGLDTR 240
DB 100 ESLQEMLVCKDIRKIYANREYHALDLTVASQAQFEVIRNGETLTGKQVTAVIDGI-- 157
QY 241 GRNDSNHEAFRGKITALY-ALGRTNANDTNGHGVAGSVLGNKXTN-----KMAPOA 295
DB 158 -----YHEDLEGRIKAFVDFVNGRERPYDNGNGTCHCADDAAAGNASISGQYRGAPRA 212
QY 296 NLVFGSMDXGGLGIGLPSNLQTLFSQAKSAGARIHTNSMGAALVNGAYTTDSRNV----- 350
DB 213 NVIGVKTILNQ-GWGSLESIMQGV-----EWCIQYNEHPPDDPIHISMSL 257
QY 351 -----DDYVR-----KNMTILPAAGNEXPNGGTISAPGTAKNAITVGATEN 392
DB 258 GGQALPYENBOEDPMVAVIVEAMNAGITVCVAGNSGPDQTLASPVSEKVTITVGLDD 317
QY 393 LRPSFGYADNINVAQFSSRGPTKGRIRKPDVMAFGTILSARSLAPDSSF-----WA 447

Db 318 -RDITREDDDD---VAPFSSRGPTTYGKPKPDILAPGVNIISLRS---PNSFYDKLQKGS 370
QY 448 NNDSTKAYMGSTSMATPIYAGNVAQUREHFVKRKGITPPPSLKALALIGA---AD----- 500
Db 371 RVGSHYTMWSTGSMATPVCGVVALMLQH---EPNLTPE--VKTRLMGTDWRMDRDN 425
QY 501 -XGLGY-----PNGNO 510
Db 426 VYCGAGTYSARGAIPNSEE 443

RESULT 4

intracellular alkaline serine proteinase aprx - Bacillus subtilis
A:Accession: A69587
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
R:Kunitz, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Allison, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entlin, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 330, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fume, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Japino, M.P.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maseel
y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A.; Authors: Schleich, S.; Schreier, R.; Scofield, P.; Sekiguchi, J.; Sekowaka, A.; Seron
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Tepstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumelein, B.; Yoshikawa, H.; Darchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MIMD:9804033; PMID:9384377
A:Accession: A69587
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-442 <KUN>
A:Cross-references: UNIPROT:O1788; UNIPARC:UPI000006043P; GB:Z99113; GB:AL009126; NID:9
A:Experimental source: strain 168
C:Genetics:
A:Gene: aprX
F:146-398/Domain: subtilisin homology <SBT>

Query Match 10.4%; Score 319.5; DB 2; Length 442;
Best Local Similarity 30.0%; Pred. No. 4.9e-15;
Matches 106; Conservative 46; Mismatches 108; Indels 93; Gaps 15;

QY 201 EYKMNDAVARGIVKA-DVAQSSVGLYGQGIIVADVADTGLDGRNDSMHEAFRGKITLY 259
Db 120 EYKALIDRTAESHAKAVRNKGITLKGAVTAAVVDGT-----YHPDLBGRK----- 168
QY 260 ALGRTNNAN-----DTNGHGHVAGSVLNGXTN-----KGMAPQANLVFQSIIMDSXGL 309
Db 169 -IGPADMVNQKTEPRDDNGHGHGHCAGDVASGSSSGQYRPAPEANLIGVAVLNQGS- 226
QY 310 GGLPSNLQTLFQSAKXAGARHITNSWGAAN-----GAYTDSBRVND 352
Db 227 GTLADIIEGV-----EKCIOYNEDNDEPIDIMSMISGIGDALYDHEQBDP 272
QY 353 YVRKND-----MTILFAAGNEKPNNGTISAPGTAKNAITVGTATLARSFQSYADNINH 406
Db 273 LVRAVEAMASAGIVCVAGNAGSGPDSQITASPEVSEKVTITVGLADNNTA-----SSDDDT 328
QY 407 VAPFSSRGPTTKDRIKPDVMAFGTILSARSSILAPDSF-----VANHDSKAYMGSTSM 461
Db 329 VAPFSSRGPTTYGKPKPDILAPGVNIISLRS---PNSFYDKLQKSSRVGSQFTMSGISM 385
QY 462 ATPYAGNVAQUREHFVKRKGITPP--PSLKALALIGAADXLGYPNNGOQ 512
Db 386 ATPICAGIALALIQ---QNPDLTPPEVKELK-----NCTDKM 420

RESULT 5

T28159
pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C:Species: Pyrococcus furiosus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
A:Accession: T28159
R:Voohorst, W.G.B.; Eggen, R.I.L.; Geejling, A.C.M.; Platteuw, C.; Sizeren, R.J.; Vos,
J. Biol. Chem. 271, 20426-20431, 1996
A:Title: Isolation and characterization of the hyperthermostable serine protease, pyroly
A:Reference number: Z20481; MIMD:96355370; PMID:8702780
A:Accession: T28159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1398 <VOO>
A:Cross-references: UNIPROT:P72186; UNIPARC:UPI000016FDS9; EMBL:U55835; NID:91556462; P
A:Experimental source: DSM3638
C:Genetics:
A:Gene: p18
C:Keywords: hydrolase; serine proteinase

Query Match 9.7%; Score 297.5; DB 2; Length 1398;
Best Local Similarity 22.9%; Pred. No. 8.5e-13;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;

QY 96 LRETKQXLEKTKAKLID-----YIPDYATVYRE-----GDYKXXXXXB 135
Db 77 LBSATTELKCLGAEILIDENRVINMLVKIKPEKXELNITSSLEKAMLRBYGLSPPIVE 136
QY 136 -HVESVPEPLPYKXIDPOLFT-----KGASXLVKAALDPTKQKKEV 176
Db 137 KVKTKKEP-----SLBPKVNSTWVYNMLQFIOEGYDSSGVAVLDTGVDPNNHFFLI 191
QY 177 QLRGLEXIAQXXXNDVAYITTKPER-KVMN-----DYARGI----- 212
Db 192 TPDGRKKIEMWDFPDGKGVDTSPFSKVNGLTIINTFYQVAGSLTNESTGLMEYVK 251
QY 213 -----YKADVAQSSGLY-----GQGQIVAVADT 236
Db 252 TYTVSNVITIGNTISANGIYHFGLLPERRYFDLNFQDQDFFYVLVNSTGNGYDIAYVDT 311
QY 237 GLDGRNDS-----SMHEAFRGKITLYALAGRTNNAN-----DTNGHGHVAG 279
Db 312 DLDYDFTBEVPLGQYNTVYDVAVFSGYTGPLNYVLAIDPNEEYAVFGDGHGHVAG 371
QY 280 SYLNGXKTN-----KGMAPQANLVFQSIIMDSXG 307
Db 372 TVAGYDSNNDADWMLSMYSGEWEVFSRLYGMDYTNVTTDVTQGVAPGAQIMAIRVLS-D 430
QY 308 GLGGLPSNLQTLFQSAKXAGARHITNSWGAANNGAYT--TDSRN--YVDYVRKNDMTILF 363
Db 431 GKGSMWDLIEGM-TYAATHGADVISMISLGG--NADYLDGTDDESVAVDBELTEKYGVVFI 487
QY 364 AAGNEKPNNGTISAPGTAKNAITVGTATLARSFQSYAD-----NTN 405
Db 488 AAGNEKPNNGTISAPGTAKNAITVGAAL-VPIINGVYYSALGYDYGYFFYFPATNV- 545
QY 406 HVAQFSSRGPTTKDRIKPDVMAFGTILSARSSILAPDSFMANHDSKAYMGSTSMATPI 465
Db 546 RIAFSSSGPRIDGKIKENNVAPAGYISSLSLPMWIGADF-----MSGTSMATPH 595
QY 466 VAGNVAQUREHFVKRKGITPPPSLKALALIGA-----DXGLGYPNNGOQMGVTLIDK 519
Db 596 VSGVALLISG-PKEBGIYVNDIITKVLVESGATVLRBDPYTGQKRTLDQHGGLVNVTK 654
QY 520 SLNVAVYVNESSXLSTSQATYXFTATACKPLKISLWSDAPASTA---SVTLVNDLIDLV 576
Db 655 SWEI-----LKAINGTTLPIVDHMDKSGSDPAEVLGVIVIGLVAR 696
QY 577 ITAPN-----GTYVGN-----DFKXPKXKNW-----DG-----RNVE----- 605
Db 697 NSIPDIVEMHIKYVDTEYRTFEIYATEPWIKPEVSGSVILLENNTBFLRVKXVDVEGLEP 756
QY 606 -----NVFINXPGSTYTIIVQAVNVVPGQKFS 634

Db 757 GLYVGRITIDPFT--TPVIEDILNTIVIPKFT 788

RESULT 6

A1341

microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #ext_change 09-Jul-2004

C:Accession: A41341; B41341; S39700; D69730

R:Stroma, A.; Ruto Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.

U: Bacteriol. 173, 6889-6895, 1991

A:Title: Cloning and characterization of the gene for an additional extracellular serine

A:Reference number: A41341; PMID:92041574; PMID:1938892

A:Accession: A41341

A:Molecule type: DNA

A:Residues: 1-806 <SLO>

A:Cross-references: UNIPROT:P29141; UNIPARC:UPI0000060C20; GB:M76590; NID:G143819; PIDN:

A:Accession: B41341

A:Molecule type: protein

A:Residues: 161-195 <SL2>

A:Cross-references: UNIPARC:UPI0000178D67

R:Glauber, P.; Kmet, F.; Arnold, M.; Coudart, M.P.; Gonzalez, W.; Hullo, M.F.; Ionescu,

., A.; Rapoport, G.; Danchin, A.

Mol. Microbiol. 10, 371-384, 1993

A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr

A:Reference number: S39655; PMID:95020537; PMID:7934828

A:Accession: S39700

A:Molecule type: DNA

A:Residues: 1-806 <GLA>

A:Cross-references: UNIPARC:UPI0000060C20; EMBL:X73124; NID:G413923; PIDN:CAA51601.1; PI

R:Kumst, F.; Ogawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Broutier, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.

Nature 390, 249-256, 1997

A:Author: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

koetter, J.; Harwood, C.R.; Hensel, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.

A:Author: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y. M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelc

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A:Author: Schleich, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Sekowaka, A.; Serot

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpeira, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wpat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K

A:Author: Yoshikawa, H.F.; Zumeeto, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; PMID:98044033; PMID:9384377

A:Accession: D69730

A:Structure: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-806 <KUN>

A:Cross-references: UNIPARC:UPI0000060C20; GB:Z99123; GB:AL009126; NID:G2636240; PIDN:CA

A:Experimental source: strain 168

C:Comment: The amino terminal sequence of the mature protein and a molecular weight of c

C:Genetics:

A:Gene: vpr

A:Start codon: TTG

C:Superfamily: microbial serine proteinase vpr; subtilisin homology

C:Keywords: hydrolase; serine proteinase

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-160/Domain: propeptide #status predicted <PRO>

F:180-548/Domain: subtilisin homology #status atypical <SBR>

Query Match 9.3%; Score 286; DB 2; Length 806;

Best local similarity 22.5%; Pred. No. 2-6e-12;

Matches 132; Conservative 60; Mismatches 188; Indels 206; Gaps 19;

Qy 155 TKGASLVKAAALDTKQKNEVQ--LRGI-----EXIAQXXXSNDVYITAKPEYKVMN 206

Db 87 TKAKKAIRK--AVKQKYNREYGVSGFSMKLPAHEIPKLAVDVAVVPMVYTKDN 144

Qy 207 DVAGIVADVAQS-----SYGLGCGQIVAVADTGLDTER----- 242

Db 145 MKDKVITISEDAVSPQMDASAFYIGANDAMDGLGYGKGIKVAIIDTGVENHPDLKKNFG 204

Qy 243 -----NDSSMHEARRGKITALYALGRNNANDTNGHTRVAGSLGNGXTYKGMAP 293

Db 205 QYKGYDVPVNDNDYDPKREPTLG-----DPRGATDHGTGVAAGV--TIKGVAP 252

Qy 294 QANLVFQSIMDSXGGLGGLPSNLQTLFQAXSAGARITNNGMGAIVNAGVATYTTDSRVNDY 353

Db 253 DATLAVRVLGRPG--STTENVIAGVRAVQDGDADVNNLSLGNLNPDMATSTAL-DW 309

Qy 354 VRKNDMTLFAAGNEXPNGGTISAPGAKNAITVGATE---NLRPSFGSY----- 400

Db 310 AMSEGVAVVTNNGNNGPMGTGSPGTSRRAISVGAQLPLNEVAVTFGSSYSAKVMGYN 369

Qy 401 -----ADNI----- 404

Db 370 KEDVYKALNKEVELVEAGIGERADFGKDLTGKVAVVKGSIAFVDKADNAKAGAIQM 429

Qy 405 -----NHVAQF 410

Db 430 VVYNNLSGEIEANVPQMSVPTIKSLBDEGKLVSKAGETKTPKLTVSKALGEQVADF 489

Qy 411 SSRGPTKD-GRKPDVMAPTXIIISARSLAPDSSFWANHDSKYAWGTSMAPTIVAGN 469

Db 490 SSRGFWMDTMKIPDISAPGVNIVSTIPTHPD-----HFYGYGSKQGTSMASPHIAGA 543

Qy 470 VAQLREHVRKRGITPKPSL--LKRAALAGAA-----DXGLGYPMNGOGMGRVTLDKSLNV 523

Db 544 VAVTIQ-----AKPKMSVEQIKALNMTVTLTKDSDEGVYPHNAQAGSARI--MNA 593

Qy 524 AYVNESSXLSTSOAKTYXFTATAGKPLKISLWSDAPASTASYTL 569

Db 594 --IKRADSLVSPGSSYGVFLKENGNETKNERFTIENGSIKSYTL 637

RESULT 7

S25835

subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain T441)

C:Species: Bacillus sp.

C>Date: 22-Nov-1993 #sequence revision 20-Feb-1995 #ext_change 05-Oct-2004

C:Accession: S25835

R:Davall, S.; Feller, G.; Narinx, B.; Gerday, C.

Gene 119, 143-144, 1992

A:Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bacillu

A:Reference number: S25835; PMID:99012966; PMID:1398082

A:Accession: S25835

A:Molecule type: DNA

A:Residues: 1-419 <DAV>

A:Cross-references: UNIPARC:UPI000008B77A; EMBL:X63533; NID:G40198; PIDN:CAA45096.1; PID

C:Superfamily: Subtilisin; subtilisin homology

C:Keywords: extracellular protein; hydrolase; serine proteinase

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-110/Domain: propeptide #status predicted <PRO>

F:111-419/Product: microbial serine proteinase #status predicted <MAT>

F:135-373/Domain: subtilisin homology <SBR>

F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 9.1%; Score 279.5; DB 1; Length 419;

Best local similarity 29.6%; Pred. No. 3e-12;

Matches 97; Conservative 47; Mismatches 133; Indels 61; Gaps 16;

Qy 171 QXNKEVOLRGLEXIAQXXXSNDVYITAKPEYKVMNVA-----RGIVKADVAQS 221

Db 81 QNNKNTLVTEKPEL-----EITATNKP-EALYNNMAASQSPWGIKALYNNSLTST 132

Qy 222 YGLYGCGQIVAVADTGLDTERNDSSMH-EAFRGKITALYALGRT--NNANDTNGHTRV 277

Db 133 SG--GAGINIAVLDTGVTNHPDLSNNVQCKD-----FTVGTFTTNSCTDRGSHGTHV 185

Qy 278 AGSVLGGXGXTNK---GMAPOANLVFQSIM--DSYGLGGLPSNLQTLFQAXSAGARITHT 332

Db 186 AGSALANGGTGSGYGVAVLPEADLWAVYLVGDGSGYADDLAERHAGDQTTALNTKTVI 245

Qy 333 N-SWGAIVNGAYTTDSRVNDYVRKNDMTILFAAGNEXPNGGTISAPGAKNAITVGATE 391

Db 246 NMSIGSGESSLIT---NAVDYAVDKGVLLIAAAGNSGPGSGIGYGAALVNAVAAALE 302
Qy 392 NLRPSGVSADNINNAVQFSSRGPTKDG-----RIKPDVMAPEGTXILSARSSLAPDSF 445
Db 303 NTLQN-GTY-----RVADFSRGRHKRYAGDYVIQKGDVEISARGAAVYST----- 346
Qy 446 WANHDSKYAVMGTSMTATPIVAGNVAQL 473
Db 347 W--FDGGYATISGISMSPHAGLAIAKT 372

RESULT 8

G83753
subtilisin-type protease (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_rev1sion 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: G83753
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, P.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: AB3650; MUID:20512582; PMID:11058132
A/Accession: G83753
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-799 <STO>
A/Cross-references: UNIPROT:Q9KEM1; UNIPARC:UPI000000C39DC; GB:AF001510; GB:BA000004; NID
A/Experimental source: strain C-125
C/Genetics:
A/Genes: vpr
C/Superfamily: microbial serine proteinase vpr; subtilisin homology
C/Keywords: hydrolase; serine proteinase
P.1-23/Domain: signal sequence #status predicted <SIG>

Query Match 9.0%; Score 276.5; DB 2; Length 799;
Best Local Similarity 23.5%; Pred. No. 1.2e-11;

Matches 149; Conservative 61; Mismatches 198; Indels 227; Gaps 25;

Qy 77 LETVPAANKLHAKQFNGP-ILR-----ETKQLEKTKGAKLIDYIPDYAVYVEYE 124
Db 50 LETV-----IVEIEDPSIIIEAKHOGKOSKRELKARQSVLEQIDLVF--SSTVTAE 100
Qy 125 GPVSPKXXIIEHVESVETLPYKXIDPOLFTKGASLVAKXALDTRQXNKVQLRIEXI 184
Db 101 YPLPSGPALE-----LPAHQIPSLIGIDVNAVYNIIEVETPDEV----- 144
Qy 185 AQXXXSNDVXYITAKPEYK-VANDVARGIVKADVAGSSYGLGQGIIVAVADTGLDTGRN 243
Db 145 -----VIEKQAVPEMLDSAPFICANDAMWAGV--TGEGLTVAIIDLTVGVYTHP 191
Qy 244 DSSMEAPFGKITALYALCRTNNANDT-----NGHTVAGSVLGNKTKMGAPQ 294
Db 192 D--LVHAF-GDYKGMDFINNDPDETPPGDPRGIEHTGTIVAGTVAANGLI-KGVAD 247
Qy 295 ANLVQSIDXSGIGLPSNLQTL--FSQAYSAGARITHTSWGAVNGAVNTTDSNVND 352
Db 248 ANLVLRVY-----GREGSGTAGVAGIRAVODGADIMLSLANTLNDPFASTAL-D 302
Qy 353 YVRKNDMTLPAAGNEXPNGGTISAPGTAKNAITVAGATENLRPSFGSY-----ADNINH 406
Db 303 WAMAGVAVLTNSNGSPNNMTVGSFGTISRDAISVQAT--RLPYKYKASVFTSDGIDY 359
Qy 407 ----- 406
Db 360 PSADIMGPSDEHLELDGEITYEYAFAGIGKPEDGEGVDEGKIALIVGRIPFVEKAE 419
Qy 407 ----- 406
Db 420 AKAAAGVAILNNVAGVQPTVGLAIPITILSNEDGLMRNELENGQNTTFSIEFDKL 479
Qy 407 -----VAQFSRSGPT-KDGRIRKPDVMAFGTXILSARSSLAPDSFMANHDSKYAVMGTSM 461
Db 480 VGEIVADFSRGPVWHMTMIKPDVSAFGVAVISTIPTHQPDPPY-----GIGSGNGTISM 533

Qy 462 ATPYVAGNVAQLREHVRKRGITPPKSLKALLIAGA-----DXGLGYENGQGWGRVTL 517
Db 534 ASPIHVAAGAAALLLEAH-PWNGV-----DHVYDALMTNAENLVDEGNRYPHNTQAG----- 584
Qy 518 DKSINAVYVNESXLSQKATY-XFTATAGKPLK 551
Db 585 --SIRIVDAIESETLVTPGSHSPGTFYTKRGQVE 617

RESULT 9

S23407
subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
C/Species: Bacillus sp.
C/Date: 04-Dec-1992 #sequence_rev1sion 04-Dec-1992 #text_change 05-Oct-2004
C/Accession: S23407
R/Marinx, B.; Davail, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A/Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarcti
A/Reference number: S23407; MUID:92256481; PMID:1581352
A/Accession: S23407
A/Molecule type: DNA
A/Residues: 1-420 <NAR>
A/Cross-references: UNIPROT:P28842; UNIPARC:UPI0000136183; EMBL:X62369; NID:940200; PID
C/Genetics:
A/Genes: sub1
C/Superfamily: Subtilisin; subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
P.1-23/Domain: signal sequence #status predicted <SIG>
P.24-111/Domain: propeptide #status predicted <PRO>
P.112-420/Product: microbial serine proteinase #status predicted <MAT>
P.136-374/Domain: subtilisin homology <SB1>
P.145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 9.0%; Score 274.5; DB 1; Length 420;
Best Local Similarity 28.3%; Pred. No. 6.9e-12;

Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

Qy 167 LDTKQXN--KEVOLRGIEIXIAQXXXSNDVXYITAKPE--YKVM---NDVARGIV---KA 215
Db 73 MKBKQFNLKKKNKLTVEKVPKL-----EATATDPEBALYNMAASQSPWGIKALYNNS 128
Qy 216 DVAQSSYGLGQGIIVAVADTGLDTRNDSSMEAPFGKITAL--YALCRT---NNANDT 270
Db 129 SITQTS---GGGGINAVLIDTVGNVTR-----HPDLRNNVBOCKQFTVGTYYTNNSCTDR 179
Qy 271 NGHGHVHVASVYLGNGKTKM---GMAPOANLVQSIM--DSXGLGLSPENLQTLPSQAS 325
Db 180 QGHGHVHVASLADGCTGNGYGVAPDADLVAYKVLGDDGSGYADIDIAAIRHAGDQATA 239
Qy 336 AGARLHTN-SWGAAVNGAYTTDSRVNDVYVRKNDMTILFPAAGNEXPNGGTISAPGTAKNA 384
Db 240 LNTKVVIMNSIGSSSESSLITNAVN---YSYKVGVLIIAAAGNSGPGYGSIGYFPAALVNA 296
Qy 385 ITVGAETNLRPSFGYADNINNAVQFSSRGPT-KDG-----RIKPDVMAPEGTXILSARSS 438
Db 297 VAAVALLEN-KYENGTY-----RVADFSRGRGYSWTDGVAIQGDVEISAPGAAYIST--- 347
Qy 439 LAPDSFMANHDSKYAVMGTSMTATPIVAGNVAQLREHVRKRGITTPPSL-----L 490
Db 348 -----W--FDGGYATISGISMSPHAGLAIAKIAQVPSASNVVDEGLYRAYENDI 398
Qy 491 KAAITAGAAD--XGLGY 505
Db 399 LSGYVAGYGDDPAGSGRPF 416

RESULT 10

S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca
N/Alternate names: subtilisin-related proteinase
C/Species: Xanthomonas campestris pv. campestris
C/Date: 21-Nov-1993 #sequence_rev1sion 07-Feb-1997 #text_change 31-Dec-2004

C:Accession: S11890
 R:Linu, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
 Mol. Gen. Genet. 220, 433-440, 1990
 A:Title: A multipurpose broad host range cloning vector and its use to characterise an
 A:Reference number: S11890; MUID:90251253; PMID:2187155
 A:Accession: S11890
 A:Molecule type: DNA
 A:Residues: 1-580 <LTI>
 A:Cross-references: UNIPROT:P23314; UNIPARC:UPI000012A398; EMBL:X51635; NID:g48533; PIDN
 A:Experimental source: Xanthomonas campestris pv. campestris
 A>Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-ALA
 C:Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
 C:Keywords: extracellular protein; hydrolase; serine proteinase
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:168-423/Domain: subtilisin homology <SBT>

Query Match 8.9%; Score 273.5; DB 2; Length 580;
 Best Local Similarity 22.7%; Pred. No. 1.3e-11;
 Matches 159; Conservative 80; Mismatches 250; Indels 211; Gaps 30;
 1 MKKKVPLSVLSAAILSTVALKNSAGAKAFDDPKIOTTTDXGFSKXQIGAAAF 60
 6 LKRRGSLTILGASALTSLMLMPAFAG-----EYLDGLATVQTHQKFTVYKDGSTAL 60
 61 LVESENVLTKXGKXKGLTVPANNGLHTXQFNGPILETKKXLEXTGAKIIDIYDVAVI 120
 61 ASPG---ALITSLRTAARAVPA-----KAGKALGLNSVRRLLAGBELV-- 100
 121 VEYEGDVAXXXXXIEHVESVEPEYLPXYIDPOLFTKGSXLVKAYALDTKQXNKEVOLRG 180
 101 -----RADRALDRAEAEFTLMRQLAADPNVSGVEV----- 129
 181 IEXTAOKXXSNDVXYITAKP-EYKVMNDVAKGIYVA--DVAOSGYLGQGIYAVADTG 237
 130 -----DQILATLTPNDTRLSEQWAFGTNAGLINIRPAMDKATGSGTVAVADTG 179
 238 L-----DTGRNDSMEAEFRGKITLALYALGRTNNAND-----TNGHGT 275
 180 ITSHADLANIAGIDFISDATTADGNGRDSNADDEGWTAAECGACIPAASSWHT 239
 276 HVAGSV--LGNGXTN-KGMAPQANLVFOSIMDSXG-----LGGLPSTL 316
 240 HVAGTVAAVTNNTTGAAGTAYGAKVVPVRLGKCGSLSDIADALVWASGGTVSGIPANA 299
 317 Q--TLFSQASAGAIHINSWGAIVNGAYTTDSRVNDVYKNDMTILPAAGNEKPNCGT 374
 300 NPAEYINNSLGGGSCSTTMON-AINGAVSNGT-----TVVVAAGNDASNVSG 346
 375 ISAPGTAKAATVGAATEN--LRPSFGYADNINHVAQSSRPGTKGRKIPVMAPGTXI 432
 347 -SLPANCANVAIVAAITTSAGAKAAYSNGTGI-----DVSAFGSI 386
 433 LSARSS--LAPDSFWANHDSKAYMGSTMATPIVAGNVAAQLREHFYKNGRGTPEK--PS 488
 387 LSTINSGTTTPGS-----ASYASVNGTSMASPHVAGVALVQS--VAPFALTPAAVET 437
 489 LLK--AALIAAAGXGLG-----PENGQGRVTLTDSLVNVAIYNBSX 531
 438 LKKTARALPGACSGCAGIVNADAATVAAIINGSGGGGGGNTLTNGTPTGTG----- 493
 532 LSTSQKATYFTATAGKPLKISLVWSDAPASTTASVTL--VNDLDVY--TAPNGTXY 585
 494 AATGELNVTITVPAG-----SGTLVTVISGSGSDLDVVRAGSAPFTDAY 539
 586 VGNDFKXPKXKNWGRNVENVFINKPQSGTYTIEVQAVN 625
 540 TCRPYRS-----GNAETCTITAP-SGYTVVRLKAYS 569

RESULT 11
 JC4908
 alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
 N:Alternate names: subtilase

C:Species: Alteromonas sp.
 C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 05-Oct-2004
 C:Accession: UC4908
 R:Tsujiido, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Inada, C.; Okami, Y.; Inamori, Y.
 Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996
 A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacte
 A:Reference number: UC4908; MUID:97141200; PMID:8987544
 A:Accession: UC4908
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-715 <TSU>
 A:Cross-references: UNIPROT:P70765; UNIPARC:UPI000017A928; DBJ:D38600; NID:g1536787; PI
 A:Experimental source: Xanthomonas campestris pv. campestris
 C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sensi
 C:Genetics:
 A:Gene: aprI
 C:Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
 C:Keywords: hydrolase
 F:1-40/Domain: signal sequence #status predicted <SIG>
 F:41-150/Domain: amino-terminal propeptide #status predicted <ATP>
 F:151-496/Product: alkaline serine protease I #status predicted <MAT>
 F:182-452/Domain: subtilisin homology <SBT>
 F:497-715/Domain: carboxy1-terminal propeptide #status predicted <CTP>
 F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.8%; Score 268.5; DB 2; Length 715;
 Best Local Similarity 25.8%; Pred. No. 3.8e-11;
 Matches 142; Conservative 51; Mismatches 201; Indels 157; Gaps 28;
 168 DTKQXNKEVQLRGLEXIKOXKXNDVXYIT-----AKREYKMD-----VARGIV 213
 117 DKLSAKGAQAFMEYVA-----SGNVEIETIDQMLKPPATNDRYNDQMYEAAAGI- 171
 214 KADVAQSSYGLYGQGIYAVADTG---LDTGRNDSMEAEFRGKITLALYALGRTNNAND 269
 172 NAPAAMDK--ATGGGVVAVALDTGVRPHLDANILPGYDMSITFVANDGARDNDARD 229
 270 -----TNG-----HGHVAG---SVGNGXKTKMGAPQANLVFQS 301
 230 PGDAVTRGECGTDSGGQVPFPAQDSSWGHVAGTVAAVTNNGEAGVAVYDAKVPVR 289
 302 IMDSXGIGLGPNSILQTLFSQASAGARIHNSWGAIV-----NGAYTTDSRVNDVY 354
 290 VL-----GKCGGLTSDIADITIASGSDRVPANANPAVYINNSLGGGACSAITQNALINQA 346
 355 RKNDMTILPAAGNEKPNCGTISAPGTAKAATVGAAT--ENLRPSFGYADNINHVAQSS 412
 347 RNNGVTIVYIAAGNDNDSNANTY-PGNCNGVNVVASVGDGSRAYSYNGANI----- 397
 413 RGPTRKGRIRKPDWMAPGTXILSARSSLAPDSFWANHDS-----KYATMGSTMATP 464
 398 -----DVAAPG---GAQSPADDEPGILSTHNSGAGPANSNDSYHSQGTSMAP 442
 465 IVAGNVAAQLR-----EHFVKN--RGITKPSILKAAI--AGAAXGLG---Y 505
 443 HVAGVAAALIKQAKSAPNDEVTILKNTTRSFASCSNGCGVDAAAAVNALGDVTP 502
 506 PNEGQGRVTLTDSLVNVAIYNBSXKLSQKATYFTATAGKPLKISLVWSDAPASTTA 565
 503 PTGN-----TLRD--GVAKTGLGGAAGSNQ--FTFEPVAGK-----TNV 538
 566 SVTL--VNDLDVITAPNGXVYVNDPFXKXKNWGR-----NNVENVFINKPQSGTYT 618
 539 TPTMSGGTGADLDLYK-----LQSQ--PTSSSDCPREYEGGNAEVCSPDAQGTTH 588
 619 IEVQAYNPVG 629
 589 VMINGYKAYSG 599

RESULT 12
 F69730
 cell wall-associated protein precursor wprA - Bacillus subtilis

C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: P69730
 R:Kumar, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choc, A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Brington, J.; Fabbre, C.; Ferrari, E. Nature 390, 249-256, 1997
 A:Authors: Poulter, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kuriita, K.; Lapidus, A.; Lardinois, A.; Luthers, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maseel, Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivaletta, C.; Rocha, B.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Schuster, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Seron, A.; Suter, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumelein, B.; Yoshikawa, H.; Danchin, A. A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: P69730
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-894 <KUN>
 A:Cross-references: UNIPROT:P54423; UNIPARC:UPI0000030283; GB:299109; GB:AL009126; NID:9
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: wprA
 F:453-664/Domain: subtilisin homology <SBR>
 Query Match 8.6%; Score 262; DB 2; Length 894;
 Best Local Similarity 23.9%; Pred. No. 1.5e-10;
 Matches 134; Conservative 92; Mismatches 211; Indels 124; Gaps 28;
 49 FSKQQTGAAP--LVSENVKXKGLKQKLETPA-----NNKLHXYQFN---G 93
 245 FDDVSENGASSYKTEKQKAIINRLYDKALQSVSFLKEBKQADRNMQLQKKTAG 304
 94 PILER-----TKQKLETKAKIIDYIPDYIVY----ETEG-----DYSKXXX 133
 305 ALITNNNTAAKSEVQIT--KVIKPYDKNSLSSVHNEMKGFSAQSKKDINVKAKKL 362
 134 IEHVESVEPYLPXYKIDPOLFTKGASXLKAXALDTKONKEVQLRGIEKXIAQ--XXXSND 192
 303 PDNLKSPF--LPKDKQKATYASAKRYKSAALITSKSNVFP---AEVQERYSLAND 416
 193 VXYITAKDEYKVNADVARGIVK--ADVA--QSSYGLYQCGQ---IYAVADTGLDGRNDS 246
 417 IQY---PYQWPLKNGGNGVGNADYKVPANTLLSKRLANDTLAVVDYDGVDSLAD-- 471
 247 MHBAKRGKITALYA--LGRTNANDTNGHGTIVAGSVLG--NGXTKGAAPQANLVFQ 300
 472 ----LKGKVRITDLGHNPFVGRNNNAMDQGGTHVAGIITAAQSDNGYSMTGLAKAKIIPV 527
 301 SIMDSXGGLGLPSNIQTLFSGQASGAGRIHTNSGAAVNGYTTDSRNVDDYVAKDMT 360
 528 KYLDSAG--SGDTGIALGICIKYADKAKAVINLSG---GGISVLEPALKYADKAVL 581
 361 ILPAAGNEXPNNGTISAPGTAKNAITVAGTENTRPSFGSYADNINHVAFSSRGPTKGR 420
 582 IAAASGNDGDN--ALSYPASSKYVMSVGT-----NRMDMTDFSYGKGL--- 625
 421 IKPDVMAAGTYILASRSLAPSSFWANHDSKRYAMGTSMATPIVAGNVALREHPYKN 480
 626 ---DLAGGSDI---PSLVPMGN-----VVTMSGTSNATPYAAAAGILFQANPGL 670
 481 RGIIFKPSILKRA--LIAGADNG-----LGPV-----NGNGMGVYTLTDSKLN 522
 671 KRTVEYDMLKTKTADISFESVDGSEELVYDGPDIETPKTPGVDMHSGYGLNWKAVS 730
 523 VA---YVNESKXLSSTQKAT 539
 731 AADLQKVNKLESTGTAVRGS 751

RESULT 13
 T29090
 Surface layer-associated STABLE protease - *Staphylococcus marinus*
 N:Alternate names: hyperthermostable protease
 C:Species: *Staphylococcus marinus*
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
 A:Accession: T29090
 R:Mayr, J.; Lupsas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J. Curr. Biol. 6, 739-749, 1996
 A:Title: A hyperthermostable protease of the subtilisin family bound to the surface layer
 A:Reference number: Z2059; MUID:96585442; PMID:8793300
 A:Accession: T29090
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1345 <MAV>
 A:Cross-references: UNIPROT:Q54437; UNIPARC:UPI0000062278; EMBL:U57968; NID:91374755; PI
 A:Experimental source: strain F1
 C:Function:
 A:Description: probably serves an exodigester function related to the organism's energy
 A:Note: sticholomeric S-layer component
 Query Match 8.5%; Score 260.5; DB 2; Length 1345;
 Best Local Similarity 29.1%; Pred. No. 3.3e-10;
 Matches 98; Conservative 41; Mismatches 105; Indels 93; Gaps 14;
 251 FKGITLALVALGRTNNANDTNGHGTIVA-----GSVL-----GNGXTNK--GMADQANLV 298
 445 YQGRYIAL-----VDFHGHGTSVAIVLASRRLVYDLGDDKLYRIMVADQAKI- 495
 299 FQSIMDSXGGLGLPSNIQTLFSGQASG-----ARHTNSW 335
 496 -----AGGDAMLLGNITLV--EAMLAGFNIIVTEEBEGYVYLSDPFGPHRADDIISNW 546
 336 GAAVNGAYTTDSRND-----DYRKXDMTILPAAGNEXPNNGTISAPGTAK 382
 547 GSTIYNFWLQCPGIDYRSSFMDELAIKRLVIGDHVITVPAAGNEXPNNGTISAPGTAK 606
 383 NAITVATE--NLRSFG--SYADNINHVAFSSRGPTKGRIRPDVMAFP-----TX 431
 607 LVITAGASTLMDYRIRIYVPGYAD--EVIKPSRGPTGQYKPDVIVNIGAFEMASR 663
 432 ILSARSLAPSSFWANHDSKRYAMGTSMATPIVAGNVALREHPVKNRGITPPPSLIK 491
 664 TIDRGYGAQPDVF-----GTSBATPPYSGTILAVFQAYKEVYMTTPDPVYAK 712
 492 AALTAGADNGIGYVNGNGWGRVTLTDSKLVAVYNE 528
 713 IILKSSAKD--IWIPIAFSGSGRDALKKADTVFPISE 747
 RESULT 14
 A33973
 high-alkaline serine proteinase (EC 3.4.21.-) YAB precursor - *Bacillus* sp. (strain YAB)
 N:Alternate names: alkaline elastase; subtilisin yab
 C:Species: *Bacillus* sp.
 C:Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 05-Oct-2004
 A:Accession: A33973
 R:Kaneko, R.; Koyama, N.; Tai, Y.C.; Huang, R.Y.; Yoda, K.; Yamasaki, M. J. Bacteriol. 171, 5232-5236, 1989
 A:Title: Molecular cloning of the structural gene for alkaline elastase yab, a new subtilisin-like serine proteinase from *Bacillus* sp. strain YAB
 A:Reference number: A33973; MUID:85359181; PMID:2670913
 A:Accession: A33973
 A:Molecule type: DNA
 A:Residues: 1-378 <KAN>
 A:Cross-references: UNIPARC:UPI0000129E78; GB:M88537; NID:9341960; PID:AAA87324.1; PID:
 A:Note: parts of this sequence, including the amino end of the mature protein, were de
 A:Accession: B33973
 A:Molecule type: protein
 A:Residues: 111-164;326-355 <KA2>
 A:Cross-references: UNIPARC:UPI0000175C84; UNIPARC:UPI0000175C85
 C:Genetics:
 A:Gene: ale

A:Start codon: TTG
C:Superfamily: Subtilisin, subtilisin homology
K:Keywords: extracellular protein, hydrolase, serine proteinase, zymogen
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-110/Domain: activation peptide #status predicted <ACP>
F:111-378/Product: alkaline elastase #status predicted <Mat>
F:123-338/Domain: subtilisin homology <SBT>
F:141,171,324/Active site: Asp, His, Ser #status predicted

Protein Sequence Searches - February 2005

All of the sequence databases on ABSs have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rnp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:33:06 ; Search time 148.883 Seconds

(without alignments)
3028.090 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059
Sequence: 1 MRKKVPLSVLSAAALSTV.....EYQATNPVPGQXSLATVN 639

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3030	99.1	639	2	Q9AOR3_9BACT
2	3022	98.8	640	2	Q9JUV9_9BACT
3	3003	98.2	640	2	Q76L84_9BACT
4	2082	68.1	433	2	Q9AOR0_9BACT
5	1952.5	63.8	434	2	Q9AOR1_9BACT
6	1948.5	63.7	433	2	Q9AOR4_9BACT
7	1941.5	63.5	433	2	Q9AOR2_9BACT
8	863.5	28.2	697	2	Q4NB18_9MICC
9	793	25.9	1748	2	Q4NVB5_9DELT
10	786.5	25.7	711	2	Q6MKR4_BDEBL
11	638	20.9	2030	2	Q747P6_GROSL
12	594	19.4	1088	2	Q4HUV5_GIBZE
13	526.5	17.2	1741	2	Q54MB8_DICDI
14	518.5	16.9	1743	1	TACG_DICDI
15	493	16.1	1825	2	Q8T9W1_DICDI
16	478	15.6	1752	2	Q9GTN7_DICDI
17	477.5	15.6	1905	1	TAGB_DICDI
18	477.5	15.6	1906	2	Q54MB3_DICDI
19	421.5	13.8	1388	2	Q560L9_9TRYP
20	413	13.5	663	2	Q501Z5_PYRKO
21	412	13.5	561	2	Q8RBZ7_THETN
22	408	13.3	654	2	Q8UBO9_PYRFU
23	391	12.8	1239	2	Q9FBA2_STRCO
24	375	12.3	430	2	Q8ENV1_OCEIH
25	368.5	12.0	1253	2	Q9RC06_STRCO
26	356	11.6	1102	2	P95684_STRCO
27	352.5	11.5	1208	2	Q82B14_STRAW
28	350.5	11.5	444	2	Q9KBJ7_BACHD
29	343.5	11.2	1139	2	Q82139_STRAW
30	340	11.1	1105	2	Q8KH66_STRVD
31	337	11.0	1245	2	Q9RL54_STRCO

ALIGNMENTS

32	336.5	11.0	412	2	Q9AER6_THERYO	Q9AER6 thermosae
33	336.5	11.0	412	2	Q8RC68_THETN	Q8RC68 thermosae
34	335	11.0	795	2	Q5NW24_9ARCH	Q5NW24 uncultured
35	333	10.9	1237	2	Q8GCT4_STRAZ	Q8GCT4 streptomyc
36	332	10.9	442	2	Q5L315_GROKA	Q5L315 geobacillus
37	332	10.9	442	2	Q651P4_BACID	Q651P4 bacillus
38	324	10.6	1294	2	Q50HM7_STRSH	Q50HM7 streptomyc
39	323	10.6	1899	2	Q67RJO_STWTH	Q67RJO symbiodacte
40	321.5	10.5	1220	2	Q910A0_STRCO	Q910A0 streptomyc
41	319.5	10.4	435	2	Q8EMJ3_OCEIH	Q8EMJ3 oceanobacti
42	319.5	10.4	442	2	Q31788_BACSU	Q31788 bacillus su
43	318.5	10.4	524	2	Q5JEB9_PYRKO	Q5JEB9 pyrococcus
44	315	10.3	818	2	Q79CG2_BACSP	Q79CG2 bacillus sp
45	315	10.3	824	2	Q45464_BACSP	Q45464 bacillus sp

RESULT 1	Q9AOR3_9BACT	PRELIMINARY;	PRT;	639 AA.
AC	Q9AOR3;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Protease.			
GN	Name=PROA;			
OS	Bacillus sp. 9860.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxId=133778;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=9860;			
RX	MEDLINE=20568675; PubMed=1118284; DOI=10.1006/brc.2000.3931;			
RA	Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,			
RA	Horiuchi K.,			
RT	"Novel oxidatively stable subtilisin-like serine proteases from			
RT	alkaliphilic Bacillus sep.: enzymatic properties, sequences, and			
RT	evolutionary relationships."			
RL	Biochem. Biophys. Res. Commun. 279:313-319(2000).			
DR	EMBL; AB046403; BAB21266.2; -; Genomic_DNA.			
DR	HSSP; P00782; 1SUP.			
DR	SMR; Q9AOR3; 206-639.			
DR	MEROPS; S08.123; -.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0004289; F:subtilase activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR007280; Pept_Bact_C.			
DR	InterPro; IPR000209; Pept_S8_S53.			
DR	Pfam; PF00082; Peptidase_S8; 1.			
DR	PRINTS; PR00723; SUBTILISIN.			
DR	PROSITE; PS00137; SUBTILASE_HIS; 1.			
DR	PROSITE; PS00138; SUBTILASE_SER; 1.			
KW	Hydrolase; Protease; Serine protease.			
SC	SEQUENCE 639 AA; 68186 MW; 316AF6FFDBE4FF54 CRC64;			
Query Match	99.1%; Score 3030; DB 2; Length 639;			
Best Local Similarity	93.4%; Pred. No. 2.5e-210; Indels 0; Gaps 0;			
Matches	597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;			
QY	1 MRKKVPLSVLSAAALSTVLSXNPSSAGAXKFDLDFGDIQTITDXXKPSKQGTGAANF	60		
DB	1 MRKKVPLSVLSAAALSTVLSXNPSSAGADATFDLDFGDIQTITDVSQFSGQRGTGAANF	60		
QY	61 LVESNVTALXGLKGLKKTTPVANNKLAHXQVNGPILBETKQXLEKTKAKIIDYIPDYVI	120		
DB	61 LVESNVTALXGLKGLKKTTPVANNKLAHXQVNGPILBETKQXLEKTKAKIIDYIPDYVI	120		
QY	121 VEYEDDVVXXXXXIEHVSVEPYLPXYXIDPOLFTKGSXLVKXALDTKQXNKEVOLRG	180		
DB	121 VEYEDDVVSKVRSIEHVSVEPYLPXYXIDPOLFTKGSXLVKXALDTKQXNKEVOLRG	180		

QY 161 IETIAQXXSNDVXYITTAKEPKYKMNVDVARGIVKADVQSSYGLYGQGOIYVADTGLDT 240
DB 161 IBEIQQVYASNDVHHITTAKEPKYKMNVDVARGIVKADVQSSYGLYGQGOIYVADTGLDT 240
QY 241 GRNDSMEHAFRGKITTAALYALGRTNNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 300
DB 241 GRNDSMEHAFRGKITTAALYALGRTNNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSQASAGARIHTNSGAAVNGAYTTDSRVNDDYRKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSQASAGARIHTNSGAAVNGAYTTDSRVNDDYRKNDMT 360
QY 361 ILFAAGNEXPNNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQFSSRGPTKQGR 420
DB 361 ILFAAGNEXPNNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQFSSRGPTKQGR 420
QY 421 IKPDVMAFGTILSARSSLAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLRHFYKX 480
DB 421 IKPDVMAFGTILSARSSLAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLRHFYKX 480
QY 481 RGITPKPSILKRALIAGAADKGLGYPNGQMGRTLLDKSLNAVAYNESSXLISTSQKATY 540
DB 481 RGITPKPSILKRALIAGAADKGLGYPNGQMGRTLLDKSLNAVAYNESSXLISTSQKATY 540
QY 541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITRAPNGTYVGNDFXXKXNDG 600
DB 541 TPTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITRAPNGTYVGNDFXXKXNDG 600
QY 601 RNVENVFINKPQSGTYTIEVOAYNVVPVGPQFSLAIYN 639
DB 601 RNVENVFINKPQSGTYTIEVOAYNVVPVGPQFSLAIYN 639

RESULT 2

Q93UV9_9BACT PRELIMINARY; PRT; 640 AA.
AC Q93UV9;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Protease.
GN Name=PROF;
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=109322;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KP43;
RA Itoh S., Saeki K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051423, BAB55674.2; -; Genomic_DNA.
DR PDB; 1WMD; X-ray; A=207-640.
DR PDB; 1WME; X-ray; A=207-640.
DR PDB; 1WMF; X-ray; A=207-640.
DR MEROPS; S08.123; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILAS_HIS; 1.
DR PROSITE; PS00138; SUBTILAS_SER; 1.
KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 640 AA; 67991 MW; 48BAF7E9D592C15 CRC64;

Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 9.6e-210;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFSLVSAABIIITVALXNPSAGARXEDLDFKGIQTTTDXGPKSKOXOTGAAL 61
DB 3 KKKKVFSLVSAABIIITVALXNPSAGARXEDLDFKGIQTTTDXGPKSKOXOTGAAL 62
QY 62 VESENVKLLKXKGLLETPANNLHLIXQFNGLLEETKQKLEXTGAKILDYIPDYAYIV 121
DB 63 VESENVKLLKXKGLLETPANNLHLIXQFNGLLEETKQKLEXTGAKILDYIPDYAYIV 122
QY 122 EYEGDVAXSKXXKIHVESVEPYLPXYXIDPOLFTKGASXLKAYALDTKQXNKEVOLRGI 181
DB 123 EYEGDVAXSSTIHHVESVEPYLPYRIDPOLFTKGASELVKAAVALDTKQXNKEVOLRGI 182
QY 182 EXIAQXXSNDVXYITTAKEPKYKMNVDVARGIVKADVQSSYGLYGQGOIYVADTGLDTG 241
DB 183 EQIAQFASINDVLYITTAKEPKYKMNVDVARGIVKADVQSSYGLYGQGOIYVADTGLDTG 242
QY 242 RNDSSMEHAFRGKITTAALYALGRTNNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 301
DB 243 RNDSSMEHAFRGKITTAALYALGRTNNANDTNGHGHVAGSVLGNKXTKGMAPQANLVFQ 302
QY 302 IMDSXGGLGGLPSNLQTLFSQASAGARIHTNSGAAVNGAYTTDSRVNDDYRKNDMTI 361
DB 303 IMDSXGGLGGLPSNLQTLFSQASAGARIHTNSGAAVNGAYTTDSRVNDDYRKNDMTI 362
QY 362 ILFAAGNEXPNNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQFSSRGPTKQGR 421
DB 363 ILFAAGNEXPNNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQFSSRGPTKQGR 422
QY 422 KPQVMAFGTILSARSSLAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLRHFYKX 481
DB 423 KPQVMAFGTILSARSSLAPDSSFMANHDSKYAYMGTSNATPIVAGNVAQLRHFYKX 482
QY 482 GITEPKPSILKRALIAGAADKGLGYPNGQMGRTLLDKSLNAVAYNESSXLISTSQKATY 541
DB 483 GITEPKPSILKRALIAGAADKGLGYPNGQMGRTLLDKSLNAVAYNESSXLISTSQKATY 542
QY 542 FTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITRAPNGTYVGNDFXXKXNDG 601
DB 543 FTATAGKPLKISLVMSDAPASTTASVTLVNDLDTITRAPNGTYVGNDFXXKXNDG 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVVPVGPQFSLAIYN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVVPVGPQFSLAIYN 640

RESULT 3

Q76L84_9BACT PRELIMINARY; PRT; 640 AA.
AC Q76L84;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Protease.
OS Bacillus sp. KSM-9865.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=192495;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-9865;
RA Okuda M., Saeki K., Kobayashi T.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB084155; BAC82522.1; -; Genomic_DNA.
DR HSP; P00782; IACN.
DR SMR; Q76L84; 207-640.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILAS_HIS; 1.
DR PROSITE; PS00137; SUBTILAS_HIS; 1.

DR PROSITE; PS00138; SUBTILASE SER; 1.
 KM Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 640 AA; 67940 MW; 421F7A150F2668F CRC64;

Query Match 98.2%; Score 3003; DB 2; Length 640;
 Best Local Similarity 92.8%; Pred. No. 2.3e-208;
 Matches 592; Conservative 1; Mismatches 45; Indels 0; Gaps 0;

QY 2 RKKKPLSLTSLAAALISTVALKLNPSAGKARXPDLPFKGIQTITDXXGFSKXQQTGAAPL 61
 DB 3 KKKKVLSTLSAAALISTVALNPSAGKARNPDPFKGIQTITDXXGFSKXQQTGAAPL 62
 QY 62 VSESNVTKLXKXGKXGKLETPANNKHIYGFNGPILEETKXQLEKXGAKLIDTIPYAYV 121
 DB 63 VSESNVTKLPKXGKXGKLETPANNKHIYGFNGPILEETKXQLEKXGAKLIDTIPYAYV 122
 QY 122 EYEGDVXSKXXXIENHVSVEPYLPYXKIDPOLFTKQASLTVAKKALDTPKXNKEVQLRGI 181
 DB 123 EYEGDVKSTSTRTEDVESVEPYLPYRIDPOLFTKQASLTVAKKALDTPKXNKEVQLRGI 182
 QY 182 EYIAQXXXSNDVYITTAKEEYVMDVARGIVKADVAQSSYGLYGGQGVAVADTGLDTG 241
 DB 183 EYIAQFATSNDVLYITTAKEEYVMDVARGIVKADVAQSSYGLYGGQGVAVADTGLDTG 242
 QY 242 RNDSSMHEAFRGKITLVALGRTNANDTNGHGTVAAGSVLGNKXTNKGMAPQANLVPOS 301
 DB 243 RNDSSMHEAFRGKITLVALGRTNANDTNGHGTVAAGSVLGNKXTNKGMAPQANLVPOS 302
 QY 302 IMDSXGGLGGLPSNLQTLFSSQAXSAGARIHTNSWGAANVAGATTDSRVNDVYRKNDMTI 361
 DB 303 IMDSXGGLGGLPSNLQTLFSSQAYSAGARIHTNSWGAANVAGATTDSRVNDVYRKNDMTI 362
 QY 362 LPAAGNEKPNGGTTISAPGTAQKAIIVGATENLRPSFGSYADININVAQPSRSGPTKDGRI 421
 DB 363 LPAAGNEKPNGGTTISAPGTAQKAIIVGATENLRPSFGSYADININVAQPSRSGPTKDGRI 422
 QY 422 KPDVVAPEGTILISARSLAPDSSFWANHDSKYAVNGGTMATPIYAGNVAOLREHFVNKR 481
 DB 423 KPDVVAPEGTILISARSLAPDSSFWANHDSKYAVNGGTMATPIYAGNVAOLREHFVNKR 482
 QY 482 GTPPESLLKALALIGAADXGLGYPNGNGMGRVTLDSLVNAYYNESSXSLSTSOKATYX 541
 DB 483 GTPPESLLKALALIGAADIGLGYNGNGMGRVTLDSLVNAYYNESSXSLSTSOKATYS 542
 QY 542 FTATAGKPLKISLVMSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFKXKXXNMDGR 601
 DB 543 FTATAGKPLKISLVMSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFKXKXXNMDGR 602
 QY 602 NNVENVPINXPOSGTYTTEVQAYNVPVGPQXESLAIYN 639
 DB 603 NNVENVPINAPQSGTYTTEVQAYNVPVGPQTESLAIYN 640

RESULT 4
 O9AOR0_9BACT PRELIMINARY; PRT; 434 AA.
 AC O9AOR0;
 DT 01-JUN-2001 (TRMBLrel. 17, Created)
 DT 01-JUN-2001 (TRMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TRMBLrel. 26, Last annotation update)
 DE Protease (Fragment).
 GN Name=PROD;
 OS Bacillus sp. NV1.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=133781;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=NV1;
 RX MEDLINE=20568675; PubMed=11118284; DOI=10.1006/dbrc.2000.3931;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 alkaliophilic Bacillus sp.: enzymatic properties, sequences, and

RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 DR EMBL; AB046406; BAB21269.1; -; Genomic_DNA.
 DR HSSP; P00782; ISUP.
 DR SMR; O9AOR0; 1-434.
 DR MEROPS; S08.123; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilisin activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR007280; Pept_Bact_C.
 DR InterPro; IPR002029; Pept_S8_S53.
 DR Pfam; PF04151; ppc; 1.
 DR Pfam; PF04151; ppc; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Protease; Serine protease.
 FT NON TER 1
 FT NON TER 1
 SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;

Query Match 68.1%; Score 2082; DB 2; Length 434;
 Best Local Similarity 91.5%; Pred. No. 4.9e-142;
 Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 206 NDVARGIKADVAQSSYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITLVALGRTN 265
 DB 1 NDVARGIKADVAQSSYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITLVALGRTN 265
 QY 266 NANDNGHGTVAAGSVLGNKXTNKGMAPQANLVPOSIMDSXGGLGGLPSNLQTLFSSQAXS 325
 DB 266 NANDNGHGTVAAGSVLGNKXTNKGMAPQANLVPOSIMDSXGGLGGLPSNLQTLFSSQAXS 325
 QY 325 AGARIHTNSWGAANVAGATTDSRVNDVYRKNDMTILPAAGNEKPNGGTTISAPGTAQKAI 385
 DB 325 AGARIHTNSWGAANVAGATTDSRVNDVYRKNDMTILPAAGNEKPNGGTTISAPGTAQKAI 385
 QY 385 LPAAGNEKPNGGTTISAPGTAQKAIIVGATENLRPSFGSYADININVAQPSRSGPTKDGRI 445
 DB 385 LPAAGNEKPNGGTTISAPGTAQKAIIVGATENLRPSFGSYADININVAQPSRSGPTKDGRI 445
 QY 445 TVGATENLRPSFGSYADININVAQPSRSGPTKDGRIKPDVVAPEGTILISARSLAPDSS 481
 DB 445 TVGATENLRPSFGSYADININVAQPSRSGPTKDGRIKPDVVAPEGTILISARSLAPDSS 481
 QY 481 WANDSKYAVNGGTMATPIYAGNVAOLREHFVNKRGITPESLLKALALIGAADXGLGY 505
 DB 481 WANDSKYAVNGGTMATPIYAGNVAOLREHFVNKRGITPESLLKALALIGAADXGLGY 505
 QY 505 PGNQGMGRVTLDSLVNAYYNESSXSLSTSOKATYXFTATAGKPLKISLVMSDAPASTTA 565
 DB 505 PGNQGMGRVTLDSLVNAYYNESSXSLSTSOKATYXFTATAGKPLKISLVMSDAPASTTA 565
 QY 565 SVTLVNDLIVITAPNGTXYVGNDFKXKXXNMDGRNNVENVPINXPOSGTYTTEVQAYN 625
 DB 565 SVTLVNDLIVITAPNGTXYVGNDFKXKXXNMDGRNNVENVPINXPOSGTYTTEVQAYN 625
 QY 625 VPVGPQXESLAIYN 639
 DB 625 VPVGPQXESLAIYN 639
 QY 639 VPGQFAPESLAIYN 434
 DB 639 VPGQFAPESLAIYN 434

RESULT 5
 O9AOR1_9BACT PRELIMINARY; PRT; 433 AA.
 AC O9AOR1;
 DT 01-JUN-2001 (TRMBLrel. 17, Created)
 DT 01-JUN-2001 (TRMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TRMBLrel. 26, Last annotation update)
 DE Protease (Fragment).
 GN Name=PROD;
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=133780;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SD521;

Medline=2056675; PubMed=1118284; DOI=10.1006/mbrc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046405; BAB21268.1; -, Genomic_DNA.
DR HSSP: Q45670; IDBI.
DR SMR: Q9AQR1; 1-433.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004289; F:subtilase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR007280; Pept_Bact_C.
DR InterPro: IPR000209; Pept_S8_S53.
DR Pfam: PF00082; Peptidase_S8; 1.
DR Pfam: PF04151; PPC; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR HydroLase; Protease; Serine protease.
FT NON_TER 1
FT TER 433
SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18FE60DDC CRC64;

Query Match 63.8%; Score 1952.5; DB 2; Length 433;

Best Local Similarity 86.6%; Pred. No. 11e-132;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

OY 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 265
DB 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 60
OY 266 NANTNGHGTIVAGSVLNGKTNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSAQXS 325
DB 61 NANTNGHGTIVAGSVLGN-ALNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSAQMN 119
OY 326 AGARHTNSWGAANVGAATTSRVNDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 120 AGARHTNSWGAANVGAATTSRVNDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 179
OY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIRKPDVMAPTXILSARSSILAPDSSF 445
DB 180 TVGATENLRPSFGSLADPNPHIAQFSSRGATRDGRIKPDVMAPTXILSARSSILAPDSSF 239
OY 446 WANHSKYAYNGGSMATPIYAGNVAQLREHFVKNRGITPKPSLKAALLAGADGXIGY 505
DB 240 WANHSKYAYNGGSMATPIYAGNVAQLREHFVKNRGITPKPSLKAALLAGADTVGIGY 299
OY 506 PNGNGGWRVTLDSKLVNVAAYNESSXLSQKATYXFTATAGKPLKISLVMSDAPASTTA 565
DB 300 PNGNGGWRVTLDSKLVNVAAYNEATALTGQKATYXFTATAGKPLKISLVMTDAPGSTTA 359
OY 566 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXXNMDGRNNVENVFINKPQSGTYTIEVOAYN 625
DB 360 SVTLVNDLDLVITAPNGKTYVGNDFSYPDNMDGRNNVENVFINKPQSGTYTIEVOAYN 419
OY 626 VPVGPOKPSLAIVN 639
DB 420 VPVGPOKPSLAIVH 433

RESULT 6

O9AQR4_9BACT PRELIMINARY; PRT; 433 AA.

AC O9AQR4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Protease (Fragment).
GN Name=PROA;
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

NCBI_TaxID=127889;
RX NCLEORTIDE SEQUENCE.

RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL: AB046402; BAB21265.1; -, Genomic_DNA.
DR HSSP: Q45670; IDBI.
DR SMR: Q9AQR4; 1-433.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004289; F:subtilase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR007280; Pept_Bact_C.
DR InterPro: IPR000209; Pept_S8_S53.
DR Pfam: PF00082; Peptidase_S8; 1.
DR Pfam: PF04151; PPC; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR HydroLase; Protease; Serine protease.
FT NON_TER 1
FT TER 433
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

Query Match 63.7%; Score 1948.5; DB 2; Length 433;

Best Local Similarity 86.4%; Pred. No. 2.2e-132;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

OY 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 265
DB 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 60
OY 266 NANTNGHGTIVAGSVLNGKTNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSAQXS 325
DB 61 NANTNGHGTIVAGSVLGN-ALNKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFSAQMN 119
OY 326 AGARHTNSWGAANVGAATTSRVNDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 120 AGARHTNSWGAANVGAATTSRVNDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 179
OY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIRKPDVMAPTXILSARSSILAPDSSF 445
DB 180 TVGATENLRPSFGSLADPNPHIAQFSSRGATRDGRIKPDVMAPTXILSARSSILAPDSSF 239
OY 446 WANHSKYAYNGGSMATPIYAGNVAQLREHFVKNRGITPKPSLKAALLAGADGXIGY 505
DB 240 WANHSKYAYNGGSMATPIYAGNVAQLREHFVKNRGITPKPSLKAALLAGADTVGIGY 299
OY 506 PNGNGGWRVTLDSKLVNVAAYNESSXLSQKATYXFTATAGKPLKISLVMSDAPASTTA 565
DB 300 PNGNGGWRVTLDSKLVNVAAYNEATALTGQKATYXFTATAGKPLKISLVMTDAPGSTTA 359
OY 566 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXXNMDGRNNVENVFINKPQSGTYTIEVOAYN 625
DB 360 SVTLVNDLDLVITAPNGKTYVGNDFSYPDNMDGRNNVENVFINKPQSGTYTIEVOAYN 419
OY 626 VPVGPOKPSLAIVN 639
DB 420 VPVGPOKPSLAIVH 433

RESULT 7

O9AQR2_9BACT PRELIMINARY; PRT; 433 AA.

AC O9AQR2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

```

DE Protease (Fragment).
GN Name=PROC;
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=133779;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RX MEDLINE=20568675; PubMed=1118284; DOI=10.1006/brc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
  Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
  alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
  evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319 (2000).
DR EMBL; AB046404; BAB21267.1; -; Genomic_DNA.
DR HSSP; Q45670; 1DBI.
DR SMR; Q9AQR2; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 433 AA; 45568 MW; B81291A803C775AE CRC64;

Query Match 63.5%; Score 1941.5; DB 2; Length 433;
Best Local Similarity 86.2%; Pred. No. 76-132;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 206 NDVAGIVADVAQSSYGYGGGQIVAVADTGLDGRNDSMHEARFGKITLALYALGRN 265
DB 1 NDVAGIVADVAQNNYGYGGQVAVADTGLDGRNDSMHEARFGKITLALYALGRN 60

QY 266 NANTNGHGTAVAGSVLGNKTKGMAPOANTVFGSINDSXGLGGLPSNLQTLFSQAXS 325
DB 61 NASDNGHGTAVAGSVLGN-ALANKGAPQANTVFGSINDSGLGLPSNLQTLFSQAMN 119

QY 326 AGARHTNSWGAUVNGAYTTDSRNVDVYRKDMTLLFPAAGNEXPGGTISAPGTAKNAI 385
DB 120 AGARHTNSWGAUVNGAYTANSRQVDEYVRNDDMTVLFAAGNEGPNSTGISAPGTAKNAI 179

QY 386 TVGATENLRPSFGSADNINNHVAQPSRSGPTQGRKPKPVMAFGTXYLSARSSLAPDSF 445
DB 180 TVGATENLRPSFGSADNPNHIAQPSRSGPTQGRKPKPVTAAGFTILSARSSLAPDSF 239

QY 446 WANHSKYAVMGSTMAATPIVAGNVAQLBEHFVKRGTGPKPSLLKAALIAAAXGAGY 505
DB 240 WANHSKYAVMGSTMAATPIVAGNVAQLBEHFVKRGTGPKPSLLKAALIAAAXGAGY 299

QY 506 PNGOGMGWVTLDKSLNVAVYNESSXLSTSQATYFTATAKPKLKLISLVSDAPASTTA 565
DB 300 PNGOGMGWVTLNKSINVAVYNEATATATGQAKATYSFGQAKPKLKLISLVMDAGSTTA 359

QY 566 STTLVNDLDTVITAPNGITXYVGNDRPKXPKKXKMGDGNNTNPFINKPQSGTTTIEVQAVN 625
DB 360 STTLVNDLDTVITAPNGQKTYGVNDPSYPYDNMWDGNNTNPFINKPQSGTYTIEVQAVN 419

QY 626 VPVGPOXPSLAIYN 639
DB 420 VPVGPOXPSLAIYH 433

```

RESULT 8
QANB18_9MICC
ID QANB18_9MICC PRELIMINARY; PRT; 697 AA.

```

AC QANB18;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, sedolisin.
GN ORFNames=ArthDRAPT_0589;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
  Micrococccineae; Micrococaceae; Arthrobacter.
OX NCBI_TaxId=290399;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
  Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24."
  Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
RL (2)
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24."
  Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
  EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
  preliminary data.
CC EMBL; AAHG0100025; EAL94539.1; -; Genomic_DNA.
CC InterPro; IPR000209; Pept_S8_S53.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 697 AA; 73821 MW; B892496C67C0714 CRC64;

```

```

Query Match 28.2%; Score 863.5; DB 2; Length 697;
Best Local Similarity 34.2%; Pred. No. 11e-53;
Matches 230; Conservative 98; Mismatches 215; Indels 129; Gaps 19;

QY 82 ANNKHI-XQFNGPILEETKQKLEXTGAKIIDYIPDYVYIESEDDYKXKXIBHVSV 140
DB 30 ASBSHILILQTAEPITTAQRAELAGIDVEMGEYSDNTYLAAPADINRRALPFGVMA 89

QY 141 EPLYEXYXIDPQLFTKGA-SKLVKXALDTQXKREVL-----RGIEIXAQXXX 189
DB 90 DVSXRVFKLPPLPRASADTGNVRSIADHBPDRRLERVDLLHPGIEAGBELIARVVA 149

QY 190 S-----NDVAYITAKREYKYMNDVYARQIVADVAQ 219
DB 150 AARVPAVAVYTPGKLRITTSVGQLELAIDEIHEIPVRROLFNNVAREIINADVQL 209

QY 220 SSYGLYGGQIVAVADTGLDGRNDSMHEARFGKITLALYALGRN--NANTNGHGTIV 277
DB 210 NGTTYRGAEEVAVADTGFDTG-DNANPHPAFTGAVQTLIYALGRAPADADPHGHGTIV 268

QY 278 AGSVLGNKXTN-----KGMAPQANTVFGSINDSXGLGGLPSNLQTLFSQAXSAGARH 331
DB 269 AGSVLGRNNSATMGGAIGTAPBALILIQSLDPRGGLGGIPVNLNDFQKTYDDGARVH 328

QY 332 TNSWGA-AVNGAYTTDSRNVDVYARK-DMTILFPAAGNEXPNG-----GTISAPGTA 381
DB 329 TNSWGVPGNLNPEYDASSREIDFVNNHDPDVCFAAGNDGVGNSDGTVDNSISGSQA 388

QY 382 KNAITVGATENLR-----SPFSY-----ADNINNHVAQPSRSGPTQGRK 422
DB 389 KNCITVGAESLSRKFTSYGTIYMPGDPFPAAPVKRDKQANPDGAVASSRSGPTGGRK 448

QY 423 PDVMAFGTYIISARSSLAP-DSSFWANHSKYAVMGSTMAATPIVAGNVAQLBEHFVNOR 481
DB 449 PDVMAFGTISLTISRNPAMGNTFGSTDPLEFFDSGTSMATPIVAGCAAVLRETLVNG 508

```

```
QY 482 GTPKPSLLKALVIA-----ADXLGYPNGNCGWRVTLDKSLNV-----AYV 526
Db 509 LNSFPAALVKALVNGADVLPQOINPSEAGBS-PNGNSGMRVNLARSVVLGGQPGNAGL 567
QY 527 NESSXLSSTOKATYYP-----TATGKPLKISLVSDAPAS 562
Db 566 GEGGPLEGQGEBSFTIDPEEVPKVAAGRRMRGPAAPALPAAGVTTLKTLVMSDDPGP 627
QY 563 TTAATVTLVNDLVLTTPANGTYTGVNDPAXKPKXXMMGGRNNVNFVFNKXPGSGTYTIVQ 622
Db 628 -----QLQNDLILVLAADGSEHGN---SGTTAGFDRNNVEQVLTMTGMPGQARIYVR 679
QY 623 AYNVVPVGPQXFS 634
Db 680 AFRTTQFPQPYA 691

RESULT 9
QANVB5_9DELT PRELIMINARY; PRT; 1748 AA.
ID QANVB5_9DELT PRELIMINARY; PRT; 1748 AA.
AC QANVB5;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, sedolisin:Nepovirus coat
protein, N-terminal:Nepovirus coat protein, N-terminal precursor.
GN ORFNames=AdenRAAT_3007;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Delta proteobacteria; Myxococcales;
OC Cyvobacteriinae; Myxococcaceae; Anaeromyxobacter.
CX NCBI_TaxId=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C."
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD0100011; EAL79523.1; -; Genomic_DNA.
KW Capsid protein; Signal.
FT SIGNAL 1 23
SQ SEQUENCE 1748 AA; 177411 MW; B1B6DE33B801DE76 CRC64;

Query Match 25.9%; Score 793; DB 2; Length 1748;
Best Local Similarity 34.9%; Pred. No. 4,4e-48;
Matches 220; Conservative 86; Mismatches 217; Indels 108; Gaps 25;

QY 88 IXQFNGPLLEETKQYLEXTGAKILRYPDYATVEYEGVXKXKXIEHVESVEPYLFX 147
Db 103 IVQSLGPRRETHKALAAAGARLGGYLBFAFVAMDETRAKVAMLFPGAVGLKRYV 162
QY 148 XIDPQLF-TKGASXLV-----KXALDTQXNKREVQJR 179
Db 163 KWAQGLADTSGAILEPGRSVLLVRVDHDLAPLVSLVRERKQKMLHVARDAASIRLL 222
QY 180 GIEKXAQXXSNDVYIT-TAKREKVMNDVARGYK----ADVASISGLIGGQGIYVA 234
Db 223 DAD-TASLAHLEEVLTMRVPALPAY-VLNDTSRMTIQYGPSPDTSISDGLGRGQIYAI 280
QY 225 DTGLD-----TGRNDSWHEAFRGKITLVALGRTNNANDTN-GHGTHVASVLGN 284
```

```
Db 281 DTGLDHACWFRPDI GAAGMHR-----KVAAYITVG--GDYDGNLGHGTHVAGTAVGD 334
QY 285 -----GXTNKGMPQANTVPOSIMDSXGLGLPSNLQTLFSQAXSAGARIHTNSWGA 338
Db 335 QTPITGGAANAGMAPGARVVVTTDLFLENNMFSPPADLAFIFTPYALGARITHNSWGS 394
QY 339 VNGAYTTDSRVVDVYKRN-DMTLLFAAGNEXPMNGGTTISAGTAKNAITVATENTLRSE 397
Db 395 SN-AVDALARADRFMEHDPFLVLFANGNAPPVGVSGAATYAKNVSVGATGN----- 448
QY 398 GSYADNINHVAFSSRGPTKGRIRKPDVMAQGYTLARSLSLAPDSFPMWHDKAYAMG 457
Db 449 GLAARD---VASFSHGPADGRKTPITLARGVIVSADSIGTAS-----NNGSTVAF-S 500
QY 458 GTSNATPIVAGNVAAQLREHFNVN---RGI-----TPKPSLLKALVIA-----AADXL 503
Db 501 GTSNATPAAAGAAALVHQYFEGGFWPSPGLGSPADARSPSALVATLVNSAQNVVAGENG 560
QY 504 G-YFNGNGMGRVTLDKSLNV-----YNESSXLSSTOKATYXFTATGKPLKISLV 555
Db 561 GPISSTQGWGRINLSNLRFPADAAYLDVVEVAAGLETGSFTROYFSTGAQPLKTLV 620
QY 556 WSDPASTTASVTLVNDLVTAPNG-TXYVGNDFX-----XPXXMMGGRNNVNF 606
Db 621 WTDAPSGQLADRSLVNDLAVTPGCAITTYLGNVFLGBSVAGAP-----DLNVBEQ 675
QY 607 VFINKPQSGTYTIEVQAYNVVPGPQXFSLA 637
Db 676 VLAAPFTVGTVTVAVTGNNVPGPQPALVI 706

RESULT 10
Q6MKR4_BDEBA PRELIMINARY; PRT; 711 AA.
ID Q6MKR4_BDEBA PRELIMINARY; PRT; 711 AA.
AC Q6MKR4;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Subtilisin-like serine protease precursor.
GN OrderedLocNames=Bd2321;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
CX NCBI_TaxId=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Bpplinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.D., Goemann A., Meyer F.;
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: 11fe cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective."
RL Science 303:689-692(2004).
DR EMBL; BX942652; CAB80143.1; -; Genomic_DNA.
DR HSBP; P27693; IAH2.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN_SER; 1.
DR PROSITE; PS00138; SUBTILAS_SER; 1.
KW Complete proteome; Protease; Signal.
FT SIGNAL 1 8
SQ SEQUENCE 711 AA; 74914 MW; 4930050843357EBS CRC64;

Query Match 25.7%; Score 786.5; DB 2; Length 711;
Best Local Similarity 30.9%; Pred. No. 4e-48;
Matches 229; Conservative 110; Mismatches 246; Indels 157; Gaps 25;

QY 8 LSVLSAAILSTVALNPNPAGXAXXFDLDFRGIGTITTDXXGFSKQXTGAAPLVESBNV 67
```

```

Db      8 ITVSTAAVLSVA-----AQAQITLKTNNAG--AIDT-----NKLTSNVAASMMMAKAT 55
Qy      68 KKKKGLKLEETVPANNKLIHQNGPPILEETKQXLEXTGAKILDYIPDYAVIVE-YEGD 126
Db      56 E-----YIVQFKKAVTEKDKALKAQ-FEVFGYLPDDALVVRGSIYS 96
Qy      127 VASXXXXIHHVSEVPEYLYKRYXD-----POLFTKASKLY-----KAXA 166
Db      97 LVTFTKTHRGVQAVVTKAGNYKNSNPASAVPTTKNNMVAVLVTFKSHVEVIAKIKR 156
Qy      167 LPTKQKKEVQ-----LRGLEXI-----AQXXSNDYXYIT 197
Db      157 MOVKVAVQVVDGHHITALLPRGLVPAVALITGVHVPAPVIRLSLFPADDEDLADYSAT 216
Qy      198 AKPEYKMDVARG--TVKDAVAQSSYGLYQGOIYAVADTGLDTRNDSMHEAFRGKI 255
Db      217 AAGDYSDLTGDSGCTLANMFDAAW-AMGYAKGQGTVMADTGLDSG-NTGAIHQDPAGGV 274
Qy      256 TALYALGRITNNA-NDTNGHGTAVAGSVLNGKTNKCM-----APQANLYFQSIIMDSKGGIG 310
Db      275 IGGYPPGLWSKMSWSDPMGHGTAVAGSVMGRTASKGLKGGAYEAMVAEGMWSPMKKYL 334
Qy      311 GLPSNLTQTFQOAXSAGARIHTNSWGA--VNGAYTTDSRVDDYVAKN-DMTILFAAGNE 368
Db      335 STPSTLGLDFEKAAPADGARIHTNSWGAATFGAIVNPVAVQVDEKSYANPDMLILFAAGNS 394
Qy      369 KP-----NGGTISAPGTAKNAITVGA TENI-----RPSFG 398
Db      395 GADKNKDRIDNSNMASTGATKKNVLTVGASENVTKSGGIQVPSIKRAAKDEWSPSPITS 454
Qy      399 ST-ANININVAOPSSSGPTKQGRIRKPDVAAFGYXILISASSLAPDSSFANHDSKAYNG 457
Db      455 STISONGNGLAFSSSGPTTDRKTEPDIVAFGTNVLSVPSQKDSPLGAAYNKQVWSG 514
Qy      458 GTSMAITPIVAGNVAQLREHFVKNRG--TPKPSILKXALLAGADKGLG----- 505
Db      515 GTSMAITPIVAGNVAQLREHFVKNRG--TPKPSILKXALLAGADKGLG----- 505
Qy      506 -----PNGNQGWGRVTLTDXSLNVA-----YNSSXSLSTSOAKATYXTATAGPLKISLV 556
Db      575 ILTRBPNSDEGVRVDVANIAMLGATQVVDNRQVAGQAGAVSYERTLMAAPSLVIANLVW 634
Qy      557 SAAAPSTTASVTLVNDLDVITAPNGTXYVGNDFPKXPKXKXNDGRNNVENVFINKPQSGT 616
Db      635 TDAPSANNAQAALVNDLDVITLTPNGQTLSPMDHI-----NNLEMIKSGILPACT 684
Qy      617 YTIIVQAVVVPV---GPOKFSL 635
Db      685 YKLTIVKGRVPOGKNGAQAVAL 706

RESULT 11
0747P6_GEOSL PRELIMINARY; PRT; 2030 AA.
ID 0747P6_GEOSL PRELIMINARY; PRT; 2030 AA.
AC 0747P6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fibronectin type III domain protein.
GN OrderedLocustNames=GSU3219;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.B., Eissen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Ward N.L., Beaman M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., Debey R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,

```

```

RA Davidson T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Ulteback T.R.,
RA Van Aken S.B., Lovley D.R., Fraser C.M.,
RA "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RA environments."
RL Science 302:1967-1969 (2003).
DR EMBL; AB017180; AAR3610.1; -; Genomic_DNA.
DR HSSP; P27693; IAH2.
DR TIGR; GSU3219; -.
DR GO; GO:0004289; P:subtilase activity; IRA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IRA.
DR InterPro; IPR011635; APHP.
DR InterPro; IPR003961; PN_III.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF07705; CARD8; 8.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILASIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
KW Complete proteome.
SQ
SEQUENCE 2030 AA; 207213 MW; ED7ADC27DD141B01 CRC64;

Query Match      20.9%; Score 638; DB 2; Length 2030;
Best Local Similarity 29.4%; Pred. No. 8.5e-37;
Matches 188; Conservative 93; Mismatches 243; Indels 116; Gaps 19;

Qy      82 ANNKLIHQNGPPILEETKQXLEXTGAKILDYIPDYAVIVEYEGDVXSKXXXIHHVSEV 141
Db      54 AIRKMWLVQFNQFVAPPEGRQLEALGCRIGDYMPTNPAVALMDDAKRAKVALLSFVEDIT 113
Qy      142 PLYLPTXIDLPOLFTGASLVKAXLD-TKQKKEVQLRG----- 181
Db      114 RPAF-----ADKLVTARKDLPAFTSIRIRIKVLAVDDPADRAAVIATLNG 161
Qy      182 -----EXIAQXXXSNDVXYITAKPEYKWNDAVARGIVKDAVQS--- 220
Db      162 NGRILNAGARTTIVPEBELAPLAQOSETAWIGVGBRLKNSDANVYQTNBYDNRTI 221
Qy      221 -SYGLYQGOIYAVADTGLD-----TGRNDSMHEAFRGKITALYALGRITNNA 270
Db      222 WEKGITGAGQIVGIDSGVDYDMPFADPNGLPGRGRIKTYG-----YDATLGDNDHVA 276
Qy      271 NGHGTAVAGSVLGN---GXTKGNAPQANLVFQSIIMDSKGLGGLPSNLQTLFQOAXSAG 327
Db      277 DGHGTHTGTTGSDGPGKPGNGIAPGARIHVDLVGTDTGLTG-SLELETYLKXAYVSG 335
Qy      328 ARHTNSGAAVNGAYTTDSRVDDYV-RXNDMTILFAAGNXPNGGTTISAPGTAKNAT 386
Db      336 ARIFNGSMGVD-SGNYDALAALDDPSWRHKDFTLAVFNNGNGPABOTRTATSPALINAT 394
Qy      387 VGATENLRPSFGSYADNINHVAAQFSSRGPTKQGRIRKPDVMAFGYXILISASSLAPDSSF 446
Db      395 VVATGN-----CTDAAT---VSABSSVQAGPADGRANPVGAPGQGVASARS---DGLLG 442
Qy      447 ANHDSKAYVMGTSMAITPIVAGNVAQLREHF-----VKNRGITPKESLQAALIA 496
Db      443 SGNSTGYMA-MSGTSSVAAAVTSGAALIRQYFDGFFPGSPAPATKQLPSAALLAVALVN 501
Qy      497 GAA-----DXGLGYNGNQGWGRVTLTDXSL-----NVAYVNSSXSLSTSOAKATYXFLTA 544
Db      502 SBAALLSDPDGSDCSKRGKMGPRGLINTLTPFNGDSHSLBVVDGGTGLSTGWLQWQRLYFS 561
Qy      545 TAGKPLKISLVSDAPASTASVTLVNDLDVITAPNGTXYVGNDF-----XPKXXKXNW 598
Db      562 PGRRLKLTITLMTDAPAPAGATSPILTNDLNVVYAPDGTTLTGNDLNSHSDYBSRTGTF 621
Qy      599 DGRNNV-ENVFINAPQSGTITIEVQAINVVPVGPQFSLAT 637
Db      622 SDRVAVBEQVAVIKPEVAGTIVLVKVGASIPVGQPPFALVM 661

RESULT 12
Q4H0T5_GIBZ8

```

Q4H0V5 GIBZE PRELIMINARY; PRT; 1088 AA.
AC Q4H0V5;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=F011223.1;
OS Glberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocrematocetidae; Hypocreales; Nectriaceae; Gibberella.
OK NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbach C., Abouelleil A., Allen N., Anderson S.,
RA Aitchison H.M., Barina N., Bastien V., Bloom T., Boguslavsky L.,
RA Boudgaltier B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lai A.,
RA Ma L.-Y., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
RA Mithova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Nordu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachtupia A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Spencer S., Schupack R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Schuster B., Strange-Thomann N., Stojanovic N., Stubbs M.,
RA Talama J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkatarman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RL "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACM01000460; EAA75433.1; -!- Genomic DNA.
CC KW Hypothetical protein.
SQ SEQUENCE 1088 AA; 119629 MW; E3B38CB94C07F542 CRC64;
Query Match 19.4%; Score 594; DB 2; Length 1088;
Best Local Similarity 31.8%; Pred. No. 5,7e-34;
Matches 197; Conservative 98; Mismatches 231; Indels 94; Gaps 29;
QY 95 ILLETKKOLEXTGAKILDIYIPYAYIVR-YEGDVAXSXXXIHHVSEVPYLPXYIDPOL 153
DB 476 ILIQKQ--LTSPYILBELKSTISLKMDWRELBRTYRVDCLIHFPNVTAKLAERI 532
QY 154 FTGKASLVKAXLADTKQXNKVEQDRIEXIAQXXXNDVYITAKPEKVMNDVARGIV 213
DB 533 AEKGVDTIK-LAVSPVRIPLTVHQDKLEALKDISRIEIV--RPD-EVLNDLARETL 588
QY 214 KADVAQSSYGLYGQCIYAVADTGLDTGRNDSM---HEAFKRTILYALGRTNAND 269
DB 589 NANITLALSTSEYGNQKVCVADTGPDOCKMDDEMILVHAFNGVHEHLALM-LGDSKD 647
QY 270 TNGHGTTHVAGSVLNGXNTN-----KGMARQANLVFOSI-----MSSXGLGSLPENT-QT 318
DB 648 TAGHGHVACASICGGGLYKNGEDIRVRGVAPEGTLTAVQVSRPNKGAIEVPMDLQ 707
QY 319 LFSQXKASAGARIHTSMGA---AVNG--AYTTDSHNVDDYV-RKNDMTILFAAGNEXENG 372
DB 708 LFSNRYKGYRIHSHSGKWDKAKTGQGLGEGQANDIDKFVIDHDPPVVLVAAGNAKA 767
QY 373 GT---IAPGTAKAIVGATENIRPSFGSYADN-----INHAQPSRSRGPTD- 418
DB 768 KSKSNHIGAAGSAFNCITVAGTTRPNNDYGFDENGAKPMWTRINDTRAKFSRSRPTKRG 827

QY 419 -----GRKPDVMAFGTYILSARS-SLAPDS-----SEFMANHDSKYAMGTSMATP 464
DB 828 RDINGNEAGRIKIKDDVAPGAILLSAASRAAKDSRRKRVWYGRGDDDMTMSGTSMT 887
QY 465 IVAGNVAQLREHFYKNGITPKPSLLAALTAGAAD-----XELGYPNGQNGRVTLD 518
DB 888 LVACCVALLREALEHKEKPSAALIRALLVNGAVNFSBQLGLGTY-DYDQGFGRVDID 946
QY 519 KSLVV-----AVYNSSSKLSSTSQ-----KATYTFAT-----AGK-PLKSLVMSD 558
DB 947 SSISMVTKLSFVDGKLPEDTQFPVAPLRQVPESEERMTSLIRVPAGRNRLVTYLAAPD 1006
QY 559 APASTASVTLVNDLIVTAPNGTYVGNDFXPXPKXXMMGRNVNVEFINXPOSQGYT 618
DB 1007 KPAQ---SGLMQNDINILVLS-GGAERHGNMGKP---GYHTNNVEKTIENENPGETFK 1059
QY 619 IEVQAY-NVPV-GPOXPSLA 636
DB 1060 IVASIMWNIDVKAFTSPAVA 1079
RESULT 13
Q54M84 D1CDI PRELIMINARY; PRT; 1741 AA.
AC Q54M84;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE ABC transporter B family protein.
GN Name=tagC; ORFNames=DOB0191192;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OK NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Bertram M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankler A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Rey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Murry D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hanser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulsged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rablinowitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Slegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.A.,
RA Williams J., Dear P.H., Noegel A.A., Barrett B., Kuspa A.;
RL "The genome of the social amoeba Dictyostelium discoideum";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAFI01000133; EAL64353.1; -!- Genomic DNA.
CC GO; GO:0016021; C:integral to membrane; IEA.
CC GO; GO:0005524; F:ATP binding; IEA.
CC GO; GO:0016887; F:ATPase activity; IEA.
CC GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
CC GO; GO:0000166; F:nucleotide binding; IEA.
CC GO; GO:0004289; F:subtilase activity; IEA.
CC GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
CC GO; GO:0006810; P:transport; IEA.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR011527; ABC_membrane_1.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transp_like.
CC InterPro; IPR000209; Pept_S8_S53.

Query Match	16.9%	Score 518.5	DB 1	Length 1743
Best Local Similarity	24.2%	Pred. No. 3e-28		
Matches 194	Conservative 114	Mismatches 213	Indels 281	Gaps 33
FT COMPBIAS 1353 1357	Poly-Asn.			
FT COMPBIAS 1358 1364	Poly-Asp.			
FT COMPBIAS 1381 1386	Poly-Asn.			
FT COMPBIAS 1707 1729	Poly-Asn.			
FT ACT_SITE 325 325	Charge relay system (By similarity)			
FT ACT_SITE 372 372	Charge relay system (By similarity)			
FT ACT_SITE 637 637	Charge relay system (By similarity)			
FT CARBOHYD 390 390	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 536 536	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 547 547	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 614 614	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 669 669	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 735 735	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 741 741	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 776 776	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 832 832	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 887 887	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 1251 1251	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 1385 1385	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 1386 1386	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 1454 1454	N-linked (GlcNAc . . .) (potential)			
FT CARBOHYD 1704 1704	N-linked (GlcNAc . . .) (potential)			
SEQ SEQUENCE 1743 AA; 194146 MW; 12DB363B2F29839 CRC64;				
Query Match	16.9%	Score 518.5	DB 1	Length 1743
Best Local Similarity	24.2%	Pred. No. 3e-28		
Matches 194	Conservative 114	Mismatches 213	Indels 281	Gaps 33
87 HIXQNGILERTKXKL-----EXTGAKILDYIFDVAIYV---EYEDVYKX 130				
115 YIVQKDRINDETRELKBEFLIGTDIVLDEQYQSHIVHYIPIHDSFLVMTQESVLSS 174				
131 XXXIHFVSVSVEPY-----LPYXIDPOLFTKGASXLVK--AXALDT--KQXN 173				
175 KEWVSWIEPESNKHILNVNEKSGIGLPYII-----LSDSTNSLQRENTLNSLTKSYN 230				
174 KEVQIRGI-----EXIAQXXNSDVXYITAKPEYK 203				
231 SKVKLTLLNQKCLKSIVYCNDSPSPSCSLINSEKLVYQWISQESN---PIERSKFKQ 287				
204 VMNDVARIIV-----KADVAQSSYGLYGQGOIVAVAVDTPLDGR--NDS----- 245				
288 TANRLSPRPVFEQTKDTLVNNDVDIP-----LRGKQILSLADTGLDHSCHPFSQSKYPI 342				
246 ---SWHEAFRGKITALVALGRTNANDNNGHGTIVAGSVTLG-----NGXTNKCMAPQA 295				
343 PLNSVNLNHR-KVYTYITTSIDSDSKYDGHGTHICGSAAGCPEDSSVNISSFSGLADA 401				
236 NLVFGSINDSGKGLGSL--PSNLQTLFSGQASAGARIHTNSNGA---AVNGAYTTDSRN 349				
402 KIAF---FDLASGSSSLTPPSDLKQLYOPLVDAGARVHCDSGVSVEGVTGSYSTPAS 458				
350 VDDYVRKN-DMTILFLAAGNEKXNGGTIS--APGTAKNALIVGATENLR-----PSGGS 399				
459 IDDFLFTHPDPFILRAAGN--NEQYLSLTLTGSTAKNIVTGAMQTIHENYITDGPYIN 515				
400 YADN----- 404				
516 YQSSVPIINOBELICDPSRYKCAVTTAQCCLESNAATTGLASCCEFTLLRKSVIDAANTQPLLY 575				
405 --NHVAQSSRGPTQDGRIKPDVMAPGTYILSARSSLA-----DSSFWMHNDKSYAM 456				
576 NENNICSPSSKQPTHDGRMKPALVAPEYITTSARSGNANTTQDCDGSIL-PVTNALLA-I 633				
457 GGTSAATPIVAGNVAQLEH-----PYKRGITPKPSLLQALALAGA----- 498				
634 SGTSAATSPALAAATTLIKQYLDVGYYPFGSIYESKKLOPFGSLLPALMINNAQLNGTFQ 693				
499 --ADKGLGYPNGN-----QGMGRVTLDKSLNVAUVNESS----- 530				
654 LITSSSITPFSQVFNENRAGASLVQGMGAKIRMSNWLHVNNNNNSNNNNKTSIDGITKFKGI 753				
531 -----XLSTSGKATYXFT-----ATAGKPLK---ISLVNSDAPASTT 564				

Db 754 GGLDLRYKPNQWKEBSLSTGNTSXCFTYKPSSSSSNSGNNIPRVATLWMTDPPSYAG 81.3

Qy 565 ASVTLVNPLDLVY-----TAPGTYXVGVGDFXXPKXXNDGRNVNENFINXP 61.2

Db 814 AKFNVLNNIDLMTMYRRDNGSTIFTSNQGSSFLG-----LAPOTDILANVEGIVANPT 86.7

Qy 613 QSGTYTIEVOAYNVVPVGPQXFS 63.4

Db 868 EPMYTRFMVAGTNVPMGPONFS 88.9

RESULT 15

Q8T9M1.DICDI PRELIMINARY, PRT, 1825 AA.

ID Q8T9M1.DICDI

AC O8T9M1

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)

DE Serine protease/ABC transporter TagD (ABC transporter B family protein).

GN Name:tagd; ORFNames=DD80191427;

OS Dictyostelium discoideum (slime mold).

OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.

OC NCBI_TaxId=44689;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Ax4;

RL Anjaard C., Loomis W.F., Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases. [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Ax4;

RL Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A., Sugaang R., Bertman M., Song J., Olsen R., Szatranski K., Xu Q., Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F., Bankier A.T., Lehman R., Hamlin N., Davies R., Gaudet P., Fey P., Pilcher K., Chen G., Saunders D., Sodergren E., Davys P., Kerhornou A., Nie X., Hall N., Anjaard C., Hemphill L., Baeson N., Fadrocher P., Desany B., Just E., Morio T., Roat R., Churcher C., Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I., Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsey R., Hauser H., James K., Qutles M., Mohan M.B., Saito T., Buchrieser C., Wardrop A., Felder M., Thangaveju M., Johnson D., Knights A., Louisse H., Mungall K., Oliver K., Price C., Quail M.A., Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M., Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivy A., Sugano S., White B., Walker D., Woodard J., Winckler T., Tanaka Y., Shalys G., Schleicher M., Wehnstock G., Rosenthal A., Cox B.C., Rasmussen R.L., Gibbs R., Loomis W.F., Platzer B., Kay R.R., Williams J., Dear P.H., Noegel A.A., Barrett B., Knapp A.; RT "The genome of the social amoeba Dictyostelium discoideum.";

RL Nature 0:0-0(2005).

EMBL AF466309; AAL/74253.1; -; Genomic_DNA.

EMBL AF0101000133; EAL64354.1; -; Genomic_DNA.

DR HSSP; P08716; 1MT0.

DR DictyBase; DD80191427; tagd.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0016887; F:ATPase activity; IEA.

DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.

DR GO; GO:0000166; F:nucleotide binding; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004289; F:subtilase activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR GO; GO:0006810; F:transport; IEA.

DR InterPro; IPR03593; AAA_ATPase.

DR InterPro; IPR011527; ABC membrane 1.

DR InterPro; IPR001140; ABC_TM_transp.

DR InterPro; IPR000209; ABC_tranp_lke.

DR Pfam; PF00604; ABC_membrane; 1.

DR Pfam: PF00082; Peptidase S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000006; ABC_transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS50929; ABC_TM1P; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1825 AA; 202642 MW; E28160BC78613A3B CRC64;

Query Match 16.1%; Score 493; DB 2; Length 1825;
Best Local Similarity 24.2%; Pred No. 2,2e-26;
Matches 192; Conservative 105; Mismatches 229; Indels 268; Gaps 31;

QY 87 HIXQFNGPILEBT---KQXLEXTG-----AKILDYIPDYAYIV--EYEGDVXSX 130
DB 127 YIVQFMDHINDEFREQKQFLINTDIVLDEQPYQSHIVYIPHDSFLVLMNDEQSNLSS 186
QY 131 XXXIYVESVER-----YLPYXIDPOLFTKGASXL-----VK 163
DB 187 KEMVSWIGFEPSPNKHILNYNEKSIGLPYIKLSDSTNSLIQRMWENTLNSIITSYNSKYK 246
QY 164 AXALPTKQNKKEVQLRGIXIAQXXX-----NDVYIITAKPEYKMMNV 208
DB 247 LTLINQKCL-KSIYCNDESSQSCSLVSEKLYQWISEQSESNYIERSEKFOYANRL 305
QY 209 ARGIV-----KADVAQSYGLYGQGIIVAVADTGLDTGR---NDS-----SM 247
DB 306 SPKALFGTMDTLVNNDRIDIP-----LRGKGILSLADTGLDGSCHFPDSNNPIPYNSV 360
QY 248 HEAFRGKITLALYALGRTNNAADTNHGHTVAGSVLGNKTN-----KMAPQANLVFO 300
DB 361 MNHRRKVVYIIGSL--HDNEDYVDGHTVCCSAAAPEDSSLAISFSGLATDAKIAEF 418
QY 301 SI-MDSXGGLGGLPBNLQTLFQSAKSAGARITNSWGA---AVNGAYTTDSRNVDDYR 355
DB 419 DIASPPSNNEPVPPEDYSQLYQPLYNAGARVHGDWSGLSIQCYLGYSYSDAGSIDDFLY 478
QY 356 KN-DWTILFAAGNEXPNGGTISAPGTAKAATVGAATENLRPSF----- 397
DB 479 THPDIILRAAGNNEYSSLS-QATAKAVITVGAEGTTHESYTTDALEYSNFEYAKST 537
QY 398 -----GSYAD-----NINHVAQ 409
DB 538 INSLCQSPDKYCTYTTAQCCTEYSTVKGLSGCCTSYKNGSYASIFSSQPELYNENNIC 597
QY 410 PSSRGPTDGRKIPDVMAFGTYILSARSLA-----PDSSFMANHDSKYAYMG 458
DB 598 PSSSKGPTHDGRKIPDIVAGQYITSARSGAATTDOCGDGLPNTVALISE-----SG 650
QY 459 TSMATPIVAGNVAOLREH-----FVKRGTIPKPSLKAALLAGADUXGLGYP-- 506
DB 651 TSMATPIATTAATTILRQYLVDSGYTPGSIYVESNKLQPTGSLKALMINNAQLNGTFPL 710
QY 507 -----NGNQGMRVTLDKSLNVAVNESS----- 530
DB 711 STNTNPSNAVPTDPAGANFVQGMGSLRMEWL---YVESSGYKPKPSRWVIGELGKDK 767
QY 531 -----XISTSQKATYXT-----ATAGKP-LKISLVMSDAPASTTASVTLVNDLD 574
DB 768 ASNWKESYLSLSTQNVSYCFTYRPSSSGSGNSGIPRIATLVWTDPPSYSGAKLNLVNNLD 827
QY 575 LVIT-----APNGTXVYG-NDPXXKXXXNMDGRNVENVF---INXPOSQTYTIE 620
DB 828 LMTNTTSEFFIYYSNGSGSYNGTKGTTPLQ---DSINNVGIGITYTPIINTYSEISFPR 884
QY 621 VQAYNVVPGPOKFS 634
DB 885 IAGTNIPIGPONFS 898

Search completed: April 7, 2006, 17:42:38
Job time : 155.883 secs

This Page Blank (uspto)

DB 241 GRDSSMHEAFRGKLTALYALGRNTNANDTNGHGHVAGSVLGNKATNKGMAPQANLVQ 300
QY 301 SIMDSXGIGLGLPSNLQTLFSQAKSAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
DB 301 SIMDSGGIGLGLPSNLQTLFSQAFSAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKAYMGTSMAPIVAGNVQOLREHFVN 480
DB 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKAYMGTSMAPIVAGNVQOLREHFVN 480
QY 481 RGITPKPSLLKRALIAGAADXLGYPNQNGMGRVTLDKSLNVAAYNESSXLSQKATY 540
DB 481 RGITPKPSLLKRALIAGAADXLGYPNQNGMGRVTLDKSLNVAAYNESSXLSQKATY 540
QY 541 XFTATAGRPKLSIYWSADAPASTTASVTLVNDLDVITAPNGTYVYGNDPFXKPKXNDG 600
DB 541 XFTATAGRPKLSIYWSADAPASTTASVTLVNDLDVITAPNGTYVYGNDPFXKPKXNDG 600
QY 601 RNNVENVFINKPOSGTYTIEVOAYNVVPGQFSLAIYN 639
DB 601 RNNVENVFINKPOSGTYTIEVOAYNVVPGQFSLAIYN 639

RESULT 2
US-09-920-954-4

; Sequence 4, Application US/0920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-920-954-4

Query Match 99.1%; Score 3030; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1,3e-258;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRKKVFLSVLSAAAILSTVALKNPISAGAXKXFDLDFKGIQTTTXXGFSKQXQTGAAP 60
DB 1 MRKKVFLSVLSAAAILSTVALKNPISAGAXKXFDLDFKGIQTTTXXGFSKQXQTGAAP 60
QY 61 LVSENVKLXKGLKLETVPANNNLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAI 120
DB 61 LVSENVKLXKGLKLETVPANNNLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAI 120
QY 121 VEYEGDVASXXXXIEHVSEVEPYLPKXIIDPOLFTFKGASXLVKALADTKQKNKEVQARG 180
DB 121 VEYEGDVASXXXXIEHVSEVEPYLPKXIIDPOLFTFKGASXLVKALADTKQKNKEVQARG 180

DB 121 VEYEGDVQSKVRSIEHVSEVEPYLPKXIIDPOLFTFKGASXLVKALADTKQKNKEVQARG 180
QY 181 IEXIAQXXXNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGGQIYVAVADTGLDT 240
DB 181 IEXIAQYASNDVHYITAKPEYKVMNDVARGIVKADVAQSSYGLYGGQIYVAVADTGLDT 240
QY 241 GRDSSMHEAFRGKLTALYALGRNTNANDTNGHGHVAGSVLGNKATNKGMAPQANLVQ 300
DB 241 GRDSSMHEAFRGKLTALYALGRNTNANDTNGHGHVAGSVLGNKATNKGMAPQANLVQ 300
QY 301 SIMDSXGIGLGLPSNLQTLFSQAKSAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
DB 301 SIMDSGGIGLGLPSNLQTLFSQAFSAGARIHTNSWGAAVNGAYTTDSRVNDVYRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKAYMGTSMAPIVAGNVQOLREHFVN 480
DB 421 IKPDVMAFGTYILSARSSSLAPDSSFMANHDSKAYMGTSMAPIVAGNVQOLREHFVN 480
QY 481 RGITPKPSLLKRALIAGAADXLGYPNQNGMGRVTLDKSLNVAAYNESSXLSQKATY 540
DB 481 RGITPKPSLLKRALIAGAADXLGYPNQNGMGRVTLDKSLNVAAYNESSXLSQKATY 540
QY 541 XFTATAGRPKLSIYWSADAPASTTASVTLVNDLDVITAPNGTYVYGNDPFXKPKXNDG 600
DB 541 XFTATAGRPKLSIYWSADAPASTTASVTLVNDLDVITAPNGTYVYGNDPFXKPKXNDG 600
QY 601 RNNVENVFINKPOSGTYTIEVOAYNVVPGQFSLAIYN 639
DB 601 RNNVENVFINKPOSGTYTIEVOAYNVVPGQFSLAIYN 639

RESULT 3
US-09-509-814A-6

; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 6,5e-258;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKVFLSVLSAAAILSTVALKNPISAGAXKXFDLDFKGIQTTTXXGFSKQXQTGAAP 61
DB 2 RKKVFLSVLSAAAILSTVALKNPISAGAXKXFDLDFKGIQTTTXXGFSKQXQTGAAP 61
QY 62 VESNVKLXKGLKLETVPANNNLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAI 121
DB 62 VESNVKLXKGLKLETVPANNNLHIXQFNGPILEETKQKLETTGAKIIDIYIPDYAI 121

```
Db 63 VSEENVKLPKGLQKLETVPANNNKHIIQFNGPILEETQOLEKTAKLIDYIPDYAYIV 122
Qy 122 EYEGDVYKXXXXIIEHVESVEPYLPYXXIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 181
Db 123 EYEGDVYKASATSTIEHVESVEPYLPYRIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 182
Qy 182 EXIAOXXXXNDVXYITTAKEPYRVMDVARGIYKADVAAOSSYGLYGGQIIVAADTGLDYG 241
Db 183 EOIAFALSNDVLYTTAKPEYRVMDVARGIYKADVAAOSSYGLYGGQIIVAADTGLDYG 242
Qy 242 RNDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPOANLVPOS 301
Db 243 RNDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPOANLVPOS 302
Qy 302 IMDSXGGLGGLPSNLTQLFSAQXVSAGARIHTNSMGAAVNGAYTTDSRNVDDYVRKNDMTI 361
Db 303 IMDSXGGLGGLPSNLTQLFSAQXVSAGARIHTNSMGAAVNGAYTTDSRNVDDYVRKNDMTI 362
Qy 362 LFAAGNEKXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDQGI 421
Db 363 LFAAGNEKXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDQGI 422
Qy 422 KPDVNAFGTXILSARSLAPDSSFMANHDSKAYMGTSMAFPYVAGNVAOLREHFVKNR 481
Db 423 KPDVNAFGTXILSARSLAPDSSFMANHDSKAYMGTSMAFPYVAGNVAOLREHFVKNR 482
Qy 482 GITPKPSLLKALLIAGAADIGLGYPNGNGMGWRTLDKSLANAYVNESSSLSTSQATYS 541
Db 483 GITPKPSLLKALLIAGAADIGLGYPNGNGMGWRTLDKSLANAYVNESSSLSTSQATYS 542
Qy 542 FTATAGKPLKISLWSDAPASTTASVTLVNDLVTITAPNGTYVGNDFXPKXNMDGR 601
Db 543 FTATAGKPLKISLWSDAPASTTASVTLVNDLVTITAPNGTYVGNDFXPKXNMDGR 602
Qy 602 NNVENVFINKPOSQTYTIEVOAYNVPGQXFSIAIVN 639
Db 603 NNVENVFINKPOSQTYTIEVOAYNVPGQXFSIAIVN 640
```

RESULT 4

```
US-09-920-954-6
; Sequence 6, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/J998/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-920-954-6
```

```
Query Match 98.8%; Score 3022; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 6; 5e-258;
```

```
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;
Qy 2 RKKKVFSLSTSAALISTALXNPSAGAKAXDDLPFGKITTTTDXKGSKQOTGAATL 61
Db 3 KKKKVFSLSTSAALISTALXNPSAGAKAXDDLPFGKITTTTDXKGSKQOTGAATL 62
Qy 62 VSEENVKLPKGLQKLETVPANNNKHIIQFNGPILEETQOLEKTAKLIDYIPDYAYIV 121
Db 63 VSEENVKLPKGLQKLETVPANNNKHIIQFNGPILEETQOLEKTAKLIDYIPDYAYIV 122
Qy 122 EYEGDVYKXXXXIIEHVESVEPYLPYXXIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 181
Db 123 EYEGDVYKASATSTIEHVESVEPYLPYRIDPOLFTKGASLYKAXALDTQXNKEVOLRGI 182
Qy 182 EXIAOXXXXNDVXYITTAKEPYRVMDVARGIYKADVAAOSSYGLYGGQIIVAADTGLDYG 241
Db 183 EOIAFALSNDVLYTTAKPEYRVMDVARGIYKADVAAOSSYGLYGGQIIVAADTGLDYG 242
Qy 242 RNDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPOANLVPOS 301
Db 243 RNDSSMHEAFRGKITLVALGRTNNANDNGHGHVAGSVLGNSTNKGMAPOANLVPOS 302
Qy 302 IMDSXGGLGGLPSNLTQLFSAQXVSAGARIHTNSMGAAVNGAYTTDSRNVDDYVRKNDMTI 361
Db 303 IMDSXGGLGGLPSNLTQLFSAQXVSAGARIHTNSMGAAVNGAYTTDSRNVDDYVRKNDMTI 362
Qy 362 LFAAGNEKXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDQGI 421
Db 363 LFAAGNEKXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTDQGI 422
Qy 422 KPDVNAFGTXILSARSLAPDSSFMANHDSKAYMGTSMAFPYVAGNVAOLREHFVKNR 481
Db 423 KPDVNAFGTXILSARSLAPDSSFMANHDSKAYMGTSMAFPYVAGNVAOLREHFVKNR 482
Qy 482 GITPKPSLLKALLIAGAADIGLGYPNGNGMGWRTLDKSLANAYVNESSSLSTSQATYS 541
Db 483 GITPKPSLLKALLIAGAADIGLGYPNGNGMGWRTLDKSLANAYVNESSSLSTSQATYS 542
Qy 542 FTATAGKPLKISLWSDAPASTTASVTLVNDLVTITAPNGTYVGNDFXPKXNMDGR 601
Db 543 FTATAGKPLKISLWSDAPASTTASVTLVNDLVTITAPNGTYVGNDFXPKXNMDGR 602
Qy 602 NNVENVFINKPOSQTYTIEVOAYNVPGQXFSIAIVN 639
Db 603 NNVENVFINKPOSQTYTIEVOAYNVPGQXFSIAIVN 640
```

RESULT 5

```
US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/J998/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
```

ORGANISM: Bacillus sp.
US-09-509-814A-8

Query Match 98.8%; Score 3021; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 8e-258;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
OY 2 RKKYFVLSAAAILSTVALXNPSAGAKAFDLDPFKGIQTTTDXGSKQQTGAAPL 61
DB 3 KKKYFVLSVLSAAAILSTVALSNPSAGARNFDLPFKGIQTTTDXGSKQQTGAAPL 62
OY 62 VESENVKLXKGLXKKLETPANNKLIHQFNGLIETKQLEKTKAKILDYIPDYAYIV 121
DB 63 VESENVKLXKGLXKKLETPANNKLIHQFNGLIETKQLEKTKAKILDYIPDYAYIV 122
OY 122 EYEGDVASXXXXIHHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKQNKKEVOLRGI 181
DB 123 EYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASELVKAVALDTKQNKKEVOLRGI 182
OY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGGIYVAVDTGLDPTG 241
DB 183 EOIAPFALSNDVLYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGGIYVAVDTGLDPTG 242
OY 242 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHTVAGSVLGNGXTNKGMAPOANLVFQS 301
DB 243 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHTVAGSVLGNGXTNKGMAPOANLVFQS 302
OY 302 IMDSXGGLGGLPSNIQTLFSGQXSAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 361
DB 303 IMDSGGLGGLPSNIQTLFSGQXSAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 362
OY 362 LPAAGNEKPNCGTTSAPGTAKNAITVGATENLRPSFGSYADININVAQFSSRGPTKGRIRI 421
DB 363 LPAAGNEKPNCGTTSAPGTAKNAITVGATENLRPSFGSYADININVAQFSSRGPTKGRIRI 422
OY 422 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYMGSTMAPIYAGNVAOLREHFVNKR 481
DB 423 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYMGSTMAPIYAGNVAOLREHFVNKR 482
OY 482 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDKSLNVAAYNNESSLSLSTQKATYS 541
DB 483 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDKSLNVAAYNNESSLSLSTQKATYS 542
OY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFPKXXKXNDGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFPKXXKXNDGR 602
OY 602 NNVENVFINKPOSGTYTIEVOAYNVPGPOKFSIAIYN 639
DB 603 NNVENVFINKPOSGTYTIEVOAYNVPGPOKFSIAIYN 640
```

RESULT 6
US-09-920-954-8
; Sequence 8, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-920-954-8

Query Match 98.8%; Score 3021; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 8e-258;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
OY 2 RKKYFVLSAAAILSTVALXNPSAGAKAFDLDPFKGIQTTTDXGSKQQTGAAPL 61
DB 3 KKKYFVLSVLSAAAILSTVALSNPSAGARNFDLPFKGIQTTTDXGSKQQTGAAPL 62
OY 62 VESENVKLXKGLXKKLETPANNKLIHQFNGLIETKQLEKTKAKILDYIPDYAYIV 121
DB 63 VESENVKLXKGLXKKLETPANNKLIHQFNGLIETKQLEKTKAKILDYIPDYAYIV 122
OY 122 EYEGDVASXXXXIHHVESVEPYLPYXIDPOLFTKGASXLVKAXALDTKQNKKEVOLRGI 181
DB 123 EYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASELVKAVALDTKQNKKEVOLRGI 182
OY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGGIYVAVDTGLDPTG 241
DB 183 EOIAPFALSNDVLYITAKPEYKVMNDVARGIVKADVAOSSYGLYGQGGIYVAVDTGLDPTG 242
OY 242 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHTVAGSVLGNGXTNKGMAPOANLVFQS 301
DB 243 RNDSSMHEAFRGKITLALYALGRTNNANDTNGHTVAGSVLGNGXTNKGMAPOANLVFQS 302
OY 302 IMDSXGGLGGLPSNIQTLFSGQXSAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 361
DB 303 IMDSGGLGGLPSNIQTLFSGQXSAGARIHTNSWGAANGAYTTDSRVVDYVRKNDMTI 362
OY 362 LPAAGNEKPNCGTTSAPGTAKNAITVGATENLRPSFGSYADININVAQFSSRGPTKGRIRI 421
DB 363 LPAAGNEKPNCGTTSAPGTAKNAITVGATENLRPSFGSYADININVAQFSSRGPTKGRIRI 422
OY 422 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYMGSTMAPIYAGNVAOLREHFVNKR 481
DB 423 KPDVMAPEGTXILSARSSLAPDSSFMANHDSKYAYMGSTMAPIYAGNVAOLREHFVNKR 482
OY 482 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDKSLNVAAYNNESSLSLSTQKATYS 541
DB 483 GITPKPSLLKALJAGAADXGLGYPNGQGWGRVTLDKSLNVAAYNNESSLSLSTQKATYS 542
OY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFPKXXKXNDGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFPKXXKXNDGR 602
OY 602 NNVENVFINKPOSGTYTIEVOAYNVPGPOKFSIAIYN 639
DB 603 NNVENVFINKPOSGTYTIEVOAYNVPGPOKFSIAIYN 640
```

RESULT 7
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE

```

FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match      98.6%; Score 3017; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 1,88-257;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 1 MRKKVPLISVLSAAAILSTVALXNPSAGARXFDLDFKGIOTTTDXXGFSKXOTGAAP 60
DB 1 MRKKVPLISVLSAAAILSTVALXNPSAGARXFDLDFKGIOTTTDXXGFSKXOTGAAP 60
QY 61 LVSESNVXLXKGLXKXKLLETVPANNKLHIQFNGPILLETQKXILEXTGAKILDIYIDYAVI 120
DB 61 LVSESNVXLXKGLXKXKLLETVPANNKLHIQFNGPILLETQKXILEXTGAKILDIYIDYAVI 120
QY 121 VEYEDDVXSKXXXIHEVSEVPEYLPYXXIDPOLFTKGASXLVKAYALDTKXNKEVQLRG 180
DB 121 VEYEDDVXSKXXXIHEVSEVPEYLPYXXIDPOLFTKGASXLVKAYALDTKXNKEVQLRG 180
QY 181 IEXIQXXXSNDVXYITAKPEYKVMNDVARGIVKADVQSSGLVGQGIIVAVADTGLDT 240
DB 181 IEXIQXXXSNDVXYITAKPEYKVMNDVARGIVKADVQSSGLVGQGIIVAVADTGLDT 240
QY 241 GRNDSMHEAFRGKITLALYALGRTNANDTNNGHTHVAGSVLGNKXTNKGMARQANLVFQ 300
DB 241 GRNDSMHEAFRGKITLALYALGRTNANDTNNGHTHVAGSVLGNKXTNKGMARQANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYERKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYERKNDMT 360
QY 361 ILFAAGNXPNGGTISAPGTAKNAITTVGATENLRPSFGSYADNINHVAQFSRGPTKDGR 420
DB 361 ILFAAGNXPNGGTISAPGTAKNAITTVGATENLRPSFGSYADNINHVAQFSRGPTKDGR 420
QY 421 IKPDVMAFGTXILSARSSSLAPDSFMANHDSKAYWGTSMATPIVAGNVALRHFVFN 480
DB 421 IKPDVMAFGTXILSARSSSLAPDSFMANHDSKAYWGTSMATPIVAGNVALRHFVFN 480
QY 481 RGITPKPSILKALTAGADXLGYPNQGWGRVTLDKSLNAVAVYNESSXILSTGKATY 540
DB 481 RGITPKPSILKALTAGADXLGYPNQGWGRVTLDKSLNAVAVYNESSXILSTGKATY 540
QY 541 XFTATAGPKIKTSLWSDAPASTTASVTLVNLDTVITPACNTXVVGNDPFXKXXXNDMG 600
DB 541 XFTATAGPKIKTSLWSDAPASTTASVTLVNLDTVITPACNTXVVGNDPFXKXXXNDMG 600
QY 601 RNNVENVPINXPOSGTYTLEVOAYNVPVGPQXFSIAIYN 639
DB 601 RNNVENVPINXPOSGTYTLEVOAYNVPVGPQXFSIAIYN 639

RESULT 8
US-09-920-954-1
; Sequence 1, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEBT, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
```

```
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
```



```
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-1
```

```
Query Match 98.6%; Score 3017; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 1.8e-257;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRKKVPLSVLSAAAILSTVALXNPSAGAKRPDIDFKGIQTTTDXKXGSKXQTGAAAF 60
DB 1 MRKKVPLSVLSAAAILSTVALXNPSAGAKRPDIDFKGIQTTTDXKXGSKXQTGAAAF 60
QY 61 LVSESNVLLKXKGLKGLSTVPANNKLHIQFNGPILEETKOKLXTGAAILDIPIDYAYI 120
DB 61 LVSESNVLLKXKGLKGLSTVPANNKLHIQFNGPILEETKOKLXTGAAILDIPIDYAYI 120
QY 121 VEYEGDVXSKXXKXIEHVESVBPYLPYXIDPOLFTKGASXLVKAXALDTKQXNKSEVOLRG 180
DB 121 VEYEGDVXSKXXKXIEHVESVBPYLPYXIDPOLFTKGASXLVKAXALDTKQXNKSEVOLRG 180
```

```
QY 181 IEXIAQXXXNDVXYITTAKEBYKMNVDVARGIVKADVQSSYGLVGGQGI VAVADTGLPT 240
DB 181 IEXIAQXXXNDVXYITTAKEBYKMNVDVARGIVKADVQSSYGLVGGQGI VAVADTGLPT 240
QY 241 GRNDSMHEAFPRGKITLALYALGRTNANDTNGHGTIVAGSVLGNXTNNGMAPOANLVFO 300
DB 241 GRNDSMHEAFPRGKITLALYALGRTNANDTNGHGTIVAGSVLGNXTNNGMAPOANLVFO 300
QY 301 SIMDSXGIGLGLPSNLQTLPSQASAGARIHTNSGAAVNGAYTTDSRRVDDYVRKNDMT 360
DB 301 SIMDSXGIGLGLPSNLQTLPSQASAGARIHTNSGAAVNGAYTTDSRRVDDYVRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPETAALNALTVAATEMLRPSFGSYADINHHVAPFSSRGPTKDR 420
DB 361 ILFAAGNEXPNGGTSAPETAALNALTVAATEMLRPSFGSYADINHHVAPFSSRGPTKDR 420
QY 421 IKPDVMAPTXILSARSSILAPDSSFMANHDSKYAYMGTSMAPIVAGVNAOLREHFVN 480
DB 421 IKPDVMAPTXILSARSSILAPDSSFMANHDSKYAYMGTSMAPIVAGVNAOLREHFVN 480
QY 481 RGITPRPSILKALIAAADXGLGYPNGNGMGVTLDKSLNAVYNESSXLSQKATY 540
DB 481 RGITPRPSILKALIAAADXGLGYPNGNGMGVTLDKSLNAVYNESSXLSQKATY 540
QY 541 XFTATAGRPDKISLYWSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFXXPKXXNMDC 600
DB 541 XFTATAGRPDKISLYWSDAPASTTASVTLVNDLIVITAPNGTXYVGNDFXXPKXXNMDC 600
QY 601 RNNVENVFINPKQSGTYTIEVQAVNVPVGPQXFSIAIYN 639
DB 601 RNNVENVFINPKQSGTYTIEVQAVNVPVGPQXFSIAIYN 639
```

```
RESULT 9
US-09-509-814A-2
Sequence 2, Application us/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIKA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASR
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
```

NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature

LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

Query Match 98.3%; Score 3007; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 1, 4e-256;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KKKVFLSVLSAAALISTVALKNPSAGXARXPDLDPKGIQTTDDKXGSKQQTGAAPFLV 62
DB 4 KKKVFLSVLSAAALISTVALKNPSAGXARXPDLDPKGIQTTDDKXGSKQQTGAAPFLV 63
QY ESENVKLXKGLXKGLLETPANNKLIHIOFNGPILEETKQKLEKGAKILDIYIPYAYIVR 122
DB 64 ESENVKLXKGLXKGLLETPANNKLIHIOFNGPILEETKQKLEKGAKILDIYIPYAYIVR 123
QY 123 YEGDVXSXXXIEHVESVEPYLPYXIDPOLFTKGASXLVPAKXLDTKQXNKEVQLGIE 182
DB 124 YEGDVXSXXXIEHVESVEPYLPYXIDPOLFTKGASXLVPAKXLDTKQXNKEVQLGIE 183
QY 183 XIAQXXXSNDVXYITTAPEYKVMNDVARGIKADVAOSSVGLYQGGQIVAVADTGLDTGR 242
DB 184 XIAQXXXSNDVXYITTAPEYKVMNDVARGIKADVAOSSVGLYQGGQIVAVADTGLDTGR 243
QY 243 NDSMHEAPFGKLTALVALGRTNANDNGHGVAGSVGNGXTNGMAPOANLVPOSI 302
DB 244 NDSMHEAPFGKLTALVALGRTNANDNGHGVAGSVGNGXTNGMAPOANLVPOSI 303
QY 303 MDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRVVDVVRKNDWTIL 362
DB 304 MDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRVVDVVRKNDWTIL 363

QY 363 PAAGNEXPNGTISAPGTANATITVGATENTLRPSFGSYADNINHVAQFSSRGPTKGRX 422
 DB 364 PAAGNEXPNGTISAPGTANATITVGATENTLRPSFGSYADNINHVAQFSSRGPTKGRX 423
 QY 423 PDVMAFGTILSARSSLPAPSSFWANHDSKYAMGTSMTPIVAGNVQLREHFVKNRG 482
 DB 424 PDVMAFGTILSARSSLPAPSSFWANHDSKYAMGTSMTPIVAGNVQLREHFVKNRG 483
 QY 483 ITPKPSLLKALLIAGAADXGLGTPNGNGQWGRITLTKSLNAVAVNNESSXLSQKATYXF 542
 DB 484 ITPKPSLLKALLIAGAADXGLGTPNGNGQWGRITLTKSLNAVAVNNESSXLSQKATYXF 543
 QY 543 TATAGPRLKISLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNDFXXKXNDGRN 602
 DB 544 TATAGPRLKISLVWSDAPASTTASVTLVNDLDTVITAPNGTXYVGNDFXXKXNDGRN 603
 QY 603 NVENVPINXPOSGTYYTIEVOAYNVVPQPOXFSIAIYN 639
 DB 604 NVENVPINXPOSGTYYTIEVOAYNVVPQPOXFSIAIYN 640
 RESULT 10
 US-09-920-954-2
 / Sequence 2, Application US/0920954
 / Patent No. 6759228
 / GENERAL INFORMATION:
 / APPLICANT: TAKAIWA, MIKIO
 / APPLICANT: OKUDA, MITSUYOSHI
 / APPLICANT: SAKI, KATSUHIISA
 / APPLICANT: KUBOTA, HIROMI
 / APPLICANT: HITOMI, JUN
 / APPLICANT: KAGEYAMA, YASUSHI
 / APPLICANT: SHIKATA, SHITSUMI
 / APPLICANT: NOMURA, MASAFUMI
 / TITLE OF INVENTION: ALKALINE PROTEASE
 / FILE REFERENCE: 0327-0832-0PCT
 / CURRENT APPLICATION NUMBER: US/09/920,954
 / PRIOR FILING DATE: 2001-08-03
 / PRIOR APPLICATION NUMBER: 09/509,814
 / PRIOR FILING DATE: 2000-04-06
 / PRIOR APPLICATION NUMBER: PCT/JP98/04528
 / PRIOR FILING DATE: 1998-10-07
 / PRIOR APPLICATION NUMBER: JP 9-274570
 / PRIOR FILING DATE: 1997-06-08
 / NUMBER OF SEQ ID NOS: 24
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 2
 / LENGTH: 640
 / TYPE: PRT
 / ORGANISM: Bacillus sp.
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (3)..(3)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (24)..(24)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (30)..(30)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (33)..(33)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (47)..(47)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (48)..(48)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (54)..(54)
 / OTHER INFORMATION: Xaa is any amino acid

/ NAME/KEY: misc_feature
 / LOCATION: (71)..(71)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (75)..(75)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (90)..(90)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (103)..(103)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (106)..(106)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (129)..(129)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (131)..(131)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (132)..(132)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (133)..(133)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (134)..(134)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (147)..(147)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (149)..(149)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (161)..(161)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (166)..(166)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (173)..(173)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (184)..(184)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (188)..(188)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (189)..(189)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (190)..(190)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (195)..(195)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (287)..(287)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (307)..(307)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (325)..(325)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature
 / LOCATION: (370)..(370)
 / OTHER INFORMATION: Xaa is any amino acid
 / NAME/KEY: misc_feature

```

1 LOCATION: (432)..(432)
2 OTHER INFORMATION: Xaa is any amino acid
3 NAME/KEY: misc_feature
4 LOCATION: (502)..(502)
5 OTHER INFORMATION: Xaa is any amino acid
6 NAME/KEY: misc_feature
7 LOCATION: (532)..(532)
8 OTHER INFORMATION: Xaa is any amino acid
9 NAME/KEY: misc_feature
10 LOCATION: (542)..(542)
11 OTHER INFORMATION: Xaa is any amino acid
12 NAME/KEY: misc_feature
13 LOCATION: (585)..(585)
14 OTHER INFORMATION: Xaa is any amino acid
15 NAME/KEY: misc_feature
16 LOCATION: (592)..(592)
17 OTHER INFORMATION: Xaa is any amino acid
18 NAME/KEY: misc_feature
19 LOCATION: (593)..(593)
20 OTHER INFORMATION: Xaa is any amino acid
21 NAME/KEY: misc_feature
22 LOCATION: (595)..(595)
23 OTHER INFORMATION: Xaa is any amino acid
24 NAME/KEY: misc_feature
25 LOCATION: (596)..(596)
26 OTHER INFORMATION: Xaa is any amino acid
27 NAME/KEY: misc_feature
28 LOCATION: (597)..(597)
29 OTHER INFORMATION: Xaa is any amino acid
30 NAME/KEY: misc_feature
31 LOCATION: (612)..(612)
32 OTHER INFORMATION: Xaa is any amino acid
33 NAME/KEY: misc_feature
34 LOCATION: (633)..(633)
35 OTHER INFORMATION: Xaa is any amino acid
36 IS-09-920-954-2

```

Query Match	98.3%	Score 3007	DB 2	Length 640
Best Local Similarity	100.0%	Pred. No. 1.4e-256		
Matches 637; Conservative	0	Mismatches 0	Indels 0	Gaps 0

Qy	3	KKKVFLSVLSAAAAIIISTVALXNPFSAGAXKFXFLDPDFKGIOTTTDDXXGSPSKOXQGTAAAAFLV	62
Db	4	KKCVFLSVLSAAAAIIISTVALXNPFSAGAXKFXFLDPDFKGIOTTTDDXXGSPSKOXQGTAAAAFLV	63
Qy	63	ESENVKLXKXKLXKKLETFVPANNKTLHXQNGPILEETKQXLEXTAKLIDTIPYAAIIVE	122
Db	64	ESENVKLXKXKLXKKLETFVPANNKTLHXQNGPILEETKQXLEXTAKLIDTIPYAAIIVE	123
Qy	123	YEGDVXSKXXXXIEHVESVEPYLPXYIIDPOLFTKQASKLVKXKALDPTKXNKEVOLRGIE	162
Db	124	YEGDVXSKXXXXIEHVESVEPYLPXYIIDPOLFTKQASKLVKXKALDPTKXNKEVOLRGIE	163
Qy	183	XIXOXKXSNVDYXITTKPEKYKWNDDVARIIVADVAQSSYGLYGGQGVVAADPGLDTR	242
Db	184	XIXOXKXSNVDYXITTKPEKYKWNDDVARIIVADVAQSSYGLYGGQGVVAADPGLDTR	243
Qy	243	NDSSMEHAFRGKITLALYALGRITNANDTIGHGTHVAGSVLGNKGXNKGMAFOANLVFQSI	302
Db	244	NDSSMEHAFRGKITLALYALGRITNANDTIGHGTHVAGSVLGNKGXNKGMAFOANLVFQSI	303
Qy	303	MDXKXGUGLGPLPSMLQTLFQOAXSAGARHTTNSWGAANVAGATTTDSRNVDPYVRKNDWTIL	362
Db	304	MDXKXGUGLGPLPSMLQTLFQOAXSAGARHTTNSWGAANVAGATTTDSRNVDPYVRKNDWTIL	363
Qy	363	FAAGNEXPNNGTISAFGTAKNAITVAGATENTLPPFSGSYADININHYAOPSSRSGPTKQRIK	442
Db	364	FAAGNEXPNNGTISAFGTAKNAITVAGATENTLPPFSGSYADININHYAOPSSRSGPTKQRIK	443
Qy	423	PDVMAPEXTILSARSSILADPSSFWANHDSKVAYMGSTSMATPIVAGANVAQLREHFVNRG	482
Db	424	PDVMAPEXTILSARSSILADPSSFWANHDSKVAYMGSTSMATPIVAGANVAQLREHFVNRG	483

[illegible]

RESULT 11
US-08-873-479-42
; Sequence 42, Application US/08873479

Query Match	89.0%	Score 2723	DB 1	Length 641
Best Local Similarity	82.9%	Pred. No. 1.5e-231		
Matches 532; Conservative	40;	Mismatches	661	Indels 4; Gaps 2

```

QY      1 MRK---KKVPLSVLSAAAILSTVALXNDPXSAGAXKFPDIDFKGIOTTTDXGPGSKOQRTGA 57
      2 1 MRKKGSKKVFSLSVLSVAALLSTVALSSPBTIGANNPBLDFKGIETTLLEKAAIKQOKTGT 60
Db      1 MRKKGSKKVFSLSVLSVAALLSTVALSSPBTIGANNPBLDFKGIETTLLEKAAIKQOKTGT 60
QY      58 AAFVLSESNVTLXKXKXKLETFVPANNKLIHXQNGFPILEETKOKLEKGAKILDTIPDY 117
      61 ASFLVNSNVKIPKSIQCKLEVPADNLTLYIVQPGPILEETQQLKEKGAKILDTIPDY 120
Db      1 ASFLVNSNVKIPKSIQCKLEVPADNLTLYIVQPGPILEETQQLKEKGAKILDTIPDY 120
QY      118 AYIVEEGDVAXXXXXIEHVSVEDEPLFYXIDLPOLFTKGASXLVAKAALDTKONKEVO 177
      121 AYIVEYDDVAVATVAIHLSEVEEYLLPYKXIDLPOLFSGAELVETVALDKOSKEVR 180

```


SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.1e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRTN 265
DB 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRTN 60
QY 266 NANTNGHGTIVAGSVLGNKXTNKGMAPQANLVFQSIIMDSGGLGGLPSNLQTLFSGQXS 325
DB 61 NANTNGHGTIVAGSVLGNKXTNKGMAPQANLVFQSIIMDSGGLGGLPSNLQTLFSGQXS 120
QY 326 AGARIHTSMGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 121 AGARIHTSMGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSLAPDSF 445
DB 181 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSLAPDSF 240
QY 446 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALIGAADVIGLY 505
DB 241 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALIGAADVIGLY 300
QY 506 PNGQGMGRVTLDSLNVAAYVNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 565
DB 301 PNGQGMGRVTLDSLNVAAYVNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 360
QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXXPKXXNMDGRNVENVFINXPOSQTYTIEVQAYN 625
DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXXPKXXNMDGRNVENVFINXPOSQTYTIEVQAYN 420
QY 626 VPVGPOKFSIAIYN 639
DB 421 VPVGPOKFSIAIYN 434

RESULT 14
US-09-985-689A-2
Sequence 2, Application US/09985689A

Patent No. 6803222
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAMA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.1e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRTN 265
DB 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRTN 60
QY 266 NANTNGHGTIVAGSVLGNKXTNKGMAPQANLVFQSIIMDSGGLGGLPSNLQTLFSGQXS 325
DB 61 NANTNGHGTIVAGSVLGNKXTNKGMAPQANLVFQSIIMDSGGLGGLPSNLQTLFSGQXS 120
QY 326 AGARIHTSMGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 385
DB 121 AGARIHTSMGAAVNGAYTTDSRNVDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 180
QY 386 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSLAPDSF 445
DB 181 TVGATENLRPSFGSYADININVAQFSSRGPTKGRIKPDVMAPTXILSARSLAPDSF 240
QY 446 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALIGAADVIGLY 505
DB 241 WANHDSKAYVGGTSMATPIVAGNVAQLREHFVNKRGITPKPSILKAALIGAADVIGLY 300
QY 506 PNGQGMGRVTLDSLNVAAYVNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 565
DB 301 PNGQGMGRVTLDSLNVAAYVNESSXLSQKATYXFTATGKPLKISLWSDAPASTTA 360
QY 566 SVTLVNDLDLVITAPNGTXYVGNDFXXPKXXNMDGRNVENVFINXPOSQTYTIEVQAYN 625
DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXXPKXXNMDGRNVENVFINXPOSQTYTIEVQAYN 420
QY 626 VPVGPOKFSIAIYN 639
DB 421 VPVGPOKFSIAIYN 434

RESULT 15
US-09-985-689A-6
Sequence 6, Application US/09985689A

Patent No. 6803222
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAMA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 6
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 68.1%; Score 2082; DB 2; Length 434;
Best Local Similarity 91.5%; Pred. No. 3e-175;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRTN 265
DB 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITALVALGRTN 60

```

Qy 266 NANDTNGHGTHTVAGSVTLGNGXTTKGMAPOANLVPOSINDSKGGLGGLPSNLTQTLFESQAXS 325
Db 61 NNDPNGHGTHTVAGSVTLGNGXTSKGMAPOANLVPOSVMDSNGGLGGLPSNVSTLFFSQAYS 120
Qy 326 AGARHTNSMGAAVNGAYTTDSRNVDYRKIDMTILPAAGNEXPNGGTISAPGTAKNAI 385
Db 121 AGARHTNSMGAAPVNGAYTTDSRNVDYRKIDMAVLPAAGNEGPNGTIISAPGTAKNAI 180
Qy 386 TVGATENTLRPSFGSYADNINHVAQFSSRGPTXGRIKPDVMAFGTYILSARSLAPDSSF 445
Db 181 TVGATENTLRPSFGSYADNINHVAQFSSRGPTXGRIKPDVMAFGTYILSARSLAPDSSF 240
Qy 446 WANNHDSKYAYMGTSMAPIVAGNVAQLREHFVQNRGITPKPSLLKALIIAGAADXGLGY 505
Db 241 WANNHDSKYAYMGTSMAPIVAGNVAQLREHFVQNRGITPKPSLLKALIIAGATDIGLGY 300
Qy 506 PGNQOQMGHVTLDKSLNVAAYNESSXLSSTOKATYXFTATAGKPLKISLVMSDAPASTTA 565
Db 301 PGNQOQMGHVTLDKSLNVAAYNBSTSLSTNOKATYSFTAQSGKPLKISLVMSDAPASTTA 360
Qy 566 SVTLVNDLIDLVTTPANGTYVGNDFPKXPKXXNMWGRNNTVENYFINXPOSQTYTIEVQAYN 625
Db 361 SVTLVNDLIDLVTTPANGTYVGNDFTPAPYDNNWMDGRNNTVENYFINAPQSQTYTIEVQAYN 420
Qy 626 VPVGPQXPSLAIYN 639
Db 421 VPQGPQAFSLAIYN 434

```

Search completed: April 7, 2006, 17:44:48
 Job time : 35.9734 secs

Y:

Y:

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using bw model

Run on: April 7, 2006, 17:58:27 ; Search time 118.407 Seconds

(without alignments)
2254.866 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059

Sequence: 1 MRKKVPLSVLSAAAILSTV.....EVOGVNVPYGPQXPSLAIYN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppaa/US10A_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3030	99.1	639	US-09-920-954-4	Sequence 4, Appl1
2	3030	99.1	639	US-10-784-870-4	Sequence 4, Appl1
3	3022	98.8	640	US-09-920-954-6	Sequence 6, Appl1
4	3022	98.8	640	US-10-456-479-4	Sequence 4, Appl1
5	3022	98.8	640	US-10-784-870-6	Sequence 6, Appl1
6	3022	98.8	640	US-10-820-712A-3	Sequence 3, Appl1
7	3022	98.8	640	US-10-820-712A-3	Sequence 3, Appl1
8	3021	98.8	640	US-09-920-954-8	Sequence 8, Appl1
9	3021	98.8	640	US-10-784-870-8	Sequence 8, Appl1
10	3017	98.6	639	US-09-920-954-1	Sequence 1, Appl1
11	3017	98.6	639	US-10-784-870-1	Sequence 1, Appl1
12	3007	98.3	640	US-09-920-954-2	Sequence 2, Appl1
13	3007	98.3	640	US-10-784-870-2	Sequence 2, Appl1
14	2155	70.4	434	US-09-985-689A-1	Sequence 1, Appl1
15	2155	70.4	434	US-09-985-689A-2	Sequence 2, Appl1
16	2155	70.4	434	US-10-456-479-2	Sequence 2, Appl1
17	2155	70.4	434	US-10-456-479-2	Sequence 2, Appl1
18	2155	70.4	434	US-10-456-479-10	Sequence 10, Appl1
19	2155	70.4	434	US-10-456-479-11	Sequence 11, Appl1
20	2155	70.4	434	US-10-837-566-1	Sequence 1, Appl1
21	2155	70.4	434	US-10-837-566-2	Sequence 2, Appl1
22	2155	70.4	434	US-10-820-712A-1	Sequence 1, Appl1
23	2155	70.4	434	US-10-820-712A-12	Sequence 12, Appl1
24	2155	70.4	434	US-10-820-714A-14	Sequence 14, Appl1
25	2155	70.4	434	US-10-820-714A-1	Sequence 1, Appl1
26	2155	70.4	434	US-10-820-714A-13	Sequence 13, Appl1
27	2155	70.4	434	US-10-820-714A-15	Sequence 15, Appl1

28	2082	68.1	434	3	US-09-985-689A-6	Sequence 6, Appl1
29	2082	68.1	434	4	US-10-456-479-15	Sequence 15, Appl1
30	2082	68.1	434	4	US-10-837-566-6	Sequence 6, Appl1
31	2082	68.1	434	5	US-10-820-712A-22	Sequence 22, Appl1
32	2082	68.1	434	5	US-10-820-714A-23	Sequence 23, Appl1
33	2060.5	67.4	433	3	US-09-985-689A-7	Sequence 7, Appl1
34	2060.5	67.4	433	4	US-10-456-479-16	Sequence 16, Appl1
35	2060.5	67.4	433	4	US-10-837-566-7	Sequence 7, Appl1
36	2060.5	67.4	433	5	US-10-820-712A-23	Sequence 23, Appl1
37	2060.5	67.4	433	5	US-10-820-714A-24	Sequence 24, Appl1
38	1952.5	63.8	433	3	US-09-985-689A-5	Sequence 5, Appl1
39	1952.5	63.8	433	4	US-10-456-479-14	Sequence 14, Appl1
40	1952.5	63.8	433	4	US-10-837-566-5	Sequence 5, Appl1
41	1952.5	63.8	433	5	US-10-820-712A-20	Sequence 20, Appl1
42	1952.5	63.8	433	5	US-10-820-714A-21	Sequence 21, Appl1
43	1948.5	63.7	433	3	US-09-985-689A-3	Sequence 3, Appl1
44	1948.5	63.7	433	4	US-10-456-479-12	Sequence 12, Appl1
45	1948.5	63.7	433	4	US-10-837-566-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1						
US-09-920-954-4						
Sequence 4, Application US/09920954						
Publication No. US20020064854A1						
GENERAL INFORMATION:						
APPLICANT: TAKAIWA, MIKIO						
APPLICANT: OKUDA, MITSUYOSHI						
APPLICANT: SAKKI, KATSUHIISA						
APPLICANT: KIBOTA, HIROMI						
APPLICANT: HIBOHI, JUN						
APPLICANT: KAGIYAMA, YASUSHI						
APPLICANT: SHIKATA, SHITSUM						
APPLICANT: NOMURA, MASAFUMI						
TITLE OF INVENTION: ALKALINE PROTEASE						
FILE REFERENCE: 0327-0832-0PCT						
CURRENT APPLICATION NUMBER: US/09/920,954						
CURRENT FILING DATE: 2001-08-03						
PRIOR APPLICATION NUMBER: 09/509,814						
PRIOR FILING DATE: 2000-04-06						
PRIOR APPLICATION NUMBER: PCT/JP98/04528						
PRIOR FILING DATE: 1998-10-07						
PRIOR APPLICATION NUMBER: JP 9-274570						
PRIOR FILING DATE: 1997-06-08						
NUMBER OF SEQ ID NOS: 24						
SOFTWARE: PatentIn version 3.0						
SEQ ID NO 4						
LENGTH: 639						
TYPE: PRT						
ORGANISM: Bacillus sp.						
US-09-920-954-4						
Query Match						
Query Local Similarity 99.1%; Score 3030; DB 3; Length 639;						
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;						
QY	1	MRKKVPLSVLSAAAILSTV	ALXNPSAGKARFDDPFGIGIOTT	TXKXGFSKXQGTGAAP	60	
DB	1	MRKKVPLSVLSAAAILSTV	ALNPNPSAGDARFDDPFGIGIOTT	TVSGFSKXQGTGAAP	60	
QY	61	LVSESNVXLKXGLKXKLTSTV	PANNLGHIXQNGPILLET	KXKXLEXTGAKIIDYIP	120	
DB	61	LVSESNVXLKXGLKXKLTSTV	PANNLGHIXQNGPILLET	KXKXLEXTGAKIIDYIP	120	
QY	121	VEYEDVYXXXXXIEHVSVEPYL	PKYKIDPDLFTKGASXLVK	KALADTKONKKEVQARG	180	
DB	121	VEYEDVYXXXXXIEHVSVEPYL	PKYKIDPDLFTKGASXLVK	KALADTKONKKEVQARG	180	
QY	181	IEIXAQQXXSNDVXYITAKPEYK	WNDVARGIVKADVQSSYGL	YGQGIYVAVADTGLDT	240	
DB	181	IEIXAQQXXSNDVXYITAKPEYK	WNDVARGIVKADVQSSYGL	YGQGIYVAVADTGLDT	240	

QY 241 GRNDSMEAFRGKITALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
DB 241 GRNDSMEAFRGKITALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSQAKSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSQAKSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMT 360
QY 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLRHEFVKN 480
DB 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLRHEFVKN 480
QY 481 RGITPKPSLLKAAALAGAADXGLGYPNGNQGMRVTLDKSLNVAAYVNESSXLSOKATY 540
DB 481 RGITPKPSLLKAAALAGAADXGLGYPNGNQGMRVTLDKSLNVAAYVNESSXLSOKATY 540
QY 541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFPXXPXXNMNDG 600
DB 541 TETATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFPXXPXXNMNDG 600
QY 601 RNNVENVFINKPQSGTYYTIEVOAYNVPGPQNFSLATVN 639
DB 601 RNNVENVFINKPQSGTYYTIEVOAYNVPGPQNFSLATVN 639

RESULT 2
US-10-784-870-4
Sequence 4, Application US/10784870
Publication No. US20040142837A1

GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT FILING DATE: 2004-02-24
PRIOR FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-4

Query Match 99.1%; Score 3030; DB 4; Length 639;
Best Local Similarity 93.4%; Pred. No. 4, 6e-261;
Matches 597; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRKKVPLSVLSAAALSTVALXNPSAGXARXFDLDFKGIQTTTDXGFSKXQGTGAALF 60
DB 1 MRKKVPLSVLSAAALSTVALXNPSAGXARXFDLDFKGIQTTTDXGFSKXQGTGAALF 60
QY 61 LVSESNVTLAKKGLKLETVPANNTLHIXQFNGPILEETKQXLEXTGAKILDYIDYAI 120
DB 61 LVSESNVTLAKKGLKLETVPANNTLHIXQFNGPILEETKQXLEXTGAKILDYIDYAI 120

QY 121 VEYEGDVASXXXXIEHVESVEPYLPXYIIDPOLFTKGSXLYKXALDTKQNKXEVQLRG 180
DB 121 VEYEGDVASKVRSLIEHVESVEPYLPKYIIDPOLFTKGSXLYKXALDTKQNKXEVQLRG 180
QY 181 IEXIAQXXSNDVXYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQGIYAAVADTGLDT 240
DB 181 IEXIAQYASNDVHYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQGIYAAVADTGLDT 240
QY 241 GRNDSMEAFRGKITALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
DB 241 GRNDSMEAFRGKITALYALGRTNANDTNGHTVAGSVLGNKTKGMAPOANLVFQ 300
QY 301 SIMDSXGGLGGLPSNLQTLFSQAKSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSQAKSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMT 360
QY 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTTISAPGTAKNAITVGATENLRPSFGSYADNINHVAFSSRGPTKGR 420
QY 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLRHEFVKN 480
DB 421 IKPDVMAFGTYILSARSSILAPDSSFWANHDSKYAVMGTSMAPIVAGNVAQLRHEFVKN 480
QY 481 RGITPKPSLLKAAALAGAADXGLGYPNGNQGMRVTLDKSLNVAAYVNESSXLSOKATY 540
DB 481 RGITPKPSLLKAAALAGAADXGLGYPNGNQGMRVTLDKSLNVAAYVNESSXLSOKATY 540
QY 541 XFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFPXXPXXNMNDG 600
DB 541 TETATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTXYVGNDFPXXPXXNMNDG 600
QY 601 RNNVENVFINKPQSGTYYTIEVOAYNVPGPQNFSLATVN 639
DB 601 RNNVENVFINKPQSGTYYTIEVOAYNVPGPQNFSLATVN 639

RESULT 3
US-09-920-954-6
Sequence 6, Application US/09920954
Publication No. US20020064854A1

GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT FILING DATE: 2001-08-03
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US/09/920, 954
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-6

Query Match 98.8%; Score 3022; DB 3; Length 640;
Best Local Similarity 93.3%; Pred. No. 2, 4e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVPLSVLSAAALSTVALXNPSAGXARXFDLDFKGIQTTTDXGFSKXQGTGAALF 61

Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

Db 3 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLPRKGIQTTTDAFGFSQGTGAAPL 62

Qy 62 VSEENVYKLLKKKKLETPANNKLIHXFNGPILEETKQXLETKAKLIDYIPDYAYIV 121

Db 63 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLPRKGIQTTTDAFGFSQGTGAAPL 62

Qy 122 EYEGDVYKXXXXIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRGI 181

Db 123 EYEGDVYKATSTIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRGI 182

Qy 182 EXIAQKXSNVYXITAKEBEYKMDVAGIYKADVAQSSYGLYGGQIYVAVADTGLDTG 241

Db 183 EYEGDVYKATSTIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRGI 182

Qy 242 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 301

Db 243 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 302

Qy 302 IMDSXGGLGGLPSNLQTLFSGAASAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDMTI 361

Db 303 IMDSXGGLGGLPSNLQTLFSGAASAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDMTI 362

Qy 362 LPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPYDGR 421

Db 363 LPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPYDGR 422

Qy 422 KPDVAPGTFTLISARSSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVAQLEHVPK 481

Db 423 KPDVAPGTFTLISARSSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVAQLEHVPK 482

Qy 482 GTPPKSLKALILAGAADXGLGYPNGNGMGRVTLDKSLNAVYNESSSLSTSQKATY 541

Db 483 GTPPKSLKALILAGAADXGLGYPNGNGMGRVTLDKSLNAVYNESSSLSTSQKATY 542

Qy 542 FTATAGKPLKISLVMSDAPASTTASVTLVNDLDTVITAPNGTXVYVGNDFXPXXNMDGR 601

Db 543 FTATAGKPLKISLVMSDAPASTTASVTLVNDLDTVITAPNGTXVYVGNDFXPXXNMDGR 602

Qy 602 NNVENVFINKPQSGTYTIEVOAYNVPGPQSFSLAIVN 639

Db 603 NNVENVFINKPQSGTYTIEVOAYNVPGPQSFSLAIVN 640

RESULT 4

US-10-456-479-4

Sequence 4, Application US/10456479

Publication No. US20040072321A1

GENERAL INFORMATION:

APPLICANT: SATO, TSUYOSHI

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: TAKIMURA, YASUSHI

APPLICANT: SUMITOMO, NOBUYUKI

APPLICANT: NOMURA, MASAFUMI

APPLICANT: KOBAYASHI, TOHICHI

TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 238700USO

CURRENT APPLICATION NUMBER: US/10/456,479

CURRENT FILING DATE: 2003-06-09

PRIOR APPLICATION NUMBER: JP 2002-186387

PRIOR FILING DATE: 2002-06-26

PRIOR APPLICATION NUMBER: JP 2002-304232

PRIOR FILING DATE: 2002-10-18

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4

LENGTH: 640

TYPE: PRT

ORGANISM: Bacillus sp. KSM-KP43

US-10-456-479-4

Query Match 98.8%; Score 3022; DB 4; Length 640;

Best Local Similarity 93.3%; Pred. No. 2.4e-260;

Common inventory
How ever there is no assigned for my appl. plan
5/25/06

Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

Qy 2 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLPRKGIQTTTDAFGFSQGTGAAPL 61

Db 3 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLPRKGIQTTTDAFGFSQGTGAAPL 62

Qy 62 VSEENVYKLLKKKKLETPANNKLIHXFNGPILEETKQXLETKAKLIDYIPDYAYIV 121

Db 63 KKKKVFSLVLSAAAIISTVALSNPSAGARNPDLPRKGIQTTTDAFGFSQGTGAAPL 62

Qy 122 EYEGDVYKXXXXIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRGI 181

Db 123 EYEGDVYKATSTIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRGI 182

Qy 182 EXIAQKXSNVYXITAKEBEYKMDVAGIYKADVAQSSYGLYGGQIYVAVADTGLDTG 241

Db 183 EYEGDVYKATSTIEHVESVEPYLPYXIDPOLFTKASLVKAXALDTQXNKEVOLRGI 182

Qy 242 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 301

Db 243 RNDSSMHEAFRGKITLVALGRTNNANDTNGHGTVAAGSVLGNKXNKMAPOANLVFOS 302

Qy 302 IMDSXGGLGGLPSNLQTLFSGAASAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDMTI 361

Db 303 IMDSXGGLGGLPSNLQTLFSGAASAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDMTI 362

Qy 362 LPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPYDGR 421

Db 363 LPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPYDGR 422

Qy 422 KPDVAPGTFTLISARSSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVAQLEHVPK 481

Db 423 KPDVAPGTFTLISARSSSLAPDSSFWANHDSKYAYMGTSMAPIVAGNVAQLEHVPK 482

Qy 482 GTPPKSLKALILAGAADXGLGYPNGNGMGRVTLDKSLNAVYNESSSLSTSQKATY 541

Db 483 GTPPKSLKALILAGAADXGLGYPNGNGMGRVTLDKSLNAVYNESSSLSTSQKATY 542

Qy 542 FTATAGKPLKISLVMSDAPASTTASVTLVNDLDTVITAPNGTXVYVGNDFXPXXNMDGR 601

Db 543 FTATAGKPLKISLVMSDAPASTTASVTLVNDLDTVITAPNGTXVYVGNDFXPXXNMDGR 602

Qy 602 NNVENVFINKPQSGTYTIEVOAYNVPGPQSFSLAIVN 639

Db 603 NNVENVFINKPQSGTYTIEVOAYNVPGPQSFSLAIVN 640

RESULT 5

US-10-784-870-6

Sequence 6, Application US/10784870

Publication No. US20040142837A1

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: SAKKI, KATSUHIISA

APPLICANT: KUBOTA, HIROMI

APPLICANT: HITOMI, JUN

APPLICANT: KAGEYAMA, YASUSHI

APPLICANT: SHIKATA, SHITSUM

APPLICANT: NOMURA, MASAFUMI

TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 0327-0832-0PCT

CURRENT APPLICATION NUMBER: US/10/784,870

CURRENT FILING DATE: 2004-02-24

PRIOR APPLICATION NUMBER: US/09/509,814A

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: PCT/JP98/04528

PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570

PRIOR FILING DATE: 1997-06-08

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn version 3.0

SEQ ID NO 6

LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-6

Query Match 98.8%; Score 3022; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,4e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
QY 2 RKKKVFLSVLSAAAILSTVALXNPSAGARXFDLDFKGIQTTDXGFSKXQGTGA AFL 61
DB 3 KKKKVFSLVLSAAAILSTVALSNPSAGARXFDLDFKGIQTTDXGFSKXQGTGA AFL 62
QY 62 VESENVKLKGLKXKLSTVPANNKLIHQFNGLILEETKQLETKAKILDIYIPDYAIV 121
DB 63 VESENVKLKPKGLQKKLETPANNKLIHQFNGLILEETKQLETKAKILDIYIPDYAIV 122
QY 122 EYEGDVXSKXXXIHEVESVEPYLPYXKIDPOLFTKGASXLVAKALDTRKXNKEVOLRGI 181
DB 123 EYEGDVKSATSTIEHVESVEPYLPYXKIDPOLFTKGASXLVAKALDTRKXNKEVOLRGI 182
QY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTG 241
DB 183 EQIAQFALSNVDVLYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTG 242
QY 242 RNDSSMHAFFRGKLTALYALGRTNNANDTNGHGVAGSVLGNGXTNKGMAPQANLVFOS 301
DB 243 RNDSSMHAFFRGKLTALYALGRTNNANDTNGHGVAGSVLGNGXTNKGMAPQANLVFOS 302
QY 302 IMDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361
DB 303 IMDSGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR I 421
DB 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR I 422
QY 422 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAVMGTSMAPIYVAGNVQQLRHFVYNR 481
DB 423 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAVMGTSMAPIYVAGNVQQLRHFVYNR 482
QY 482 GITPKPSLLKALALAGAADXGLGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 541
DB 483 GITPKPSLLKALALAGAADXGLGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFXXPKXXNM DGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFXXPKXXNM DGR 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVPGPQFSLAIYN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVPGPQFSLAIYN 640
```

RESULT 6
US-10-820-712A-3
Sequence 3, Application US/10820712A
Publication No. US20050026804A1

GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Toru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820,712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent version 3.2

Search Report

SEQ ID NO 3
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp. KSM-KP43
US-10-820-712A-3

Query Match 98.8%; Score 3022; DB 5; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,4e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

```
QY 2 RKKKVFLSVLSAAAILSTVALXNPSAGARXFDLDFKGIQTTDXGFSKXQGTGA AFL 61
DB 3 KKKKVFSLVLSAAAILSTVALSNPSAGARXFDLDFKGIQTTDXGFSKXQGTGA AFL 62
QY 62 VESENVKLKGLKXKLSTVPANNKLIHQFNGLILEETKQLETKAKILDIYIPDYAIV 121
DB 63 VESENVKLKPKGLQKKLETPANNKLIHQFNGLILEETKQLETKAKILDIYIPDYAIV 122
QY 122 EYEGDVXSKXXXIHEVESVEPYLPYXKIDPOLFTKGASXLVAKALDTRKXNKEVOLRGI 181
DB 123 EYEGDVKSATSTIEHVESVEPYLPYXKIDPOLFTKGASXLVAKALDTRKXNKEVOLRGI 182
QY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTG 241
DB 183 EQIAQFALSNVDVLYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTG 242
QY 242 RNDSSMHAFFRGKLTALYALGRTNNANDTNGHGVAGSVLGNGXTNKGMAPQANLVFOS 301
DB 243 RNDSSMHAFFRGKLTALYALGRTNNANDTNGHGVAGSVLGNGXTNKGMAPQANLVFOS 302
QY 302 IMDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 361
DB 303 IMDSGGLGGLPSNLQTLFSGQXSAGARIHTNSMGAAVNGAYTTDSRVNDVYRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR I 421
DB 363 LFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQFSSRGPTKGR I 422
QY 422 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAVMGTSMAPIYVAGNVQQLRHFVYNR 481
DB 423 KPDVMAFGTXILSARSSLAPDSSFWANHDSKYAVMGTSMAPIYVAGNVQQLRHFVYNR 482
QY 482 GITPKPSLLKALALAGAADXGLGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 541
DB 483 GITPKPSLLKALALAGAADXGLGYPNGOGMGRVTLDSLVNAVYNESSXLSSTOKATYX 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFXXPKXXNM DGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDVITAPNGTYVGNDFXXPKXXNM DGR 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVPGPQFSLAIYN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVPGPQFSLAIYN 640
```

RESULT 7
US-10-820-714A-3
Sequence 3, Application US/10820714A
Publication No. US20050214922A1

GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Kobayashi, Toru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasuhiro
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697USO
CURRENT APPLICATION NUMBER: US/10/820,714A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 24

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 640
; TYPE: PR
; ORGANISM: Bacillus sp. KSM-KP43
US-10-820-714A-3

Query Match
Best Local Similarity 98.8%; Score 3022; DB 5; Length 640;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFSLVLSAAAILSTVALKNPSAGXARFPLDFKGIQTTTDXKGFSGKXQOTGAAPL 61
DB 3 KKKKVFSLVLSAAAILSTVALSNPSAGARNPDLDFKGIQTTTDAKGFSGKXQOTGAAPL 62
QY 62 VSEENVKLGKGLKKLETPANNKLIHQPNPILBETKQLEXTGAKILDYIPDYAYIV 121
DB 63 VSEENVKLGKGLKKLETPANNKLIHQPNPILBETKQLEXTGAKILDYIPDYAYIV 122
QY 122 EYEGDVSKXXXXIEHVESVEPYLPYXIDPOLFTKGASLVKAXALDTQXNKEVOLRGI 181
DB 123 EYEGDVSKATSTIEHVESVEPYLPYRIDPOLFTKGASLVKAVALDTQXNKEVOLRGI 182
QY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIVAADTGLDGT 241
DB 183 EQLAQPAISNDVLYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIVAADTGLDGT 242
QY 242 RNDSSWHEAFRGKITLVALGRTNNANDTNGHTHVASVLGNKGXTNKGMAPOANLVFQS 301
DB 243 RNDSSWHEAFRGKITLVALGRTNNANDTNGHTHVASVLGNKGXTNKGMAPOANLVFQS 302
QY 302 IMDSXGGLGGLPSNQTLPQXASAGARIHTNSWGAANVGAAYTTDSRNVDYVRKNDMTI 361
DB 303 IMDSXGGLGGLPSNQTLPQASAGARIHTNSWGAANVGAAYTTDSRNVDYVRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGAKNAITVGATENLRPSGSYADNINHVAQSSSRGPTDGR 421
DB 363 LFAAGNEXPNGGTISAPGAKNAITVGATENLRPSGSYADNINHVAQSSSRGPTDGR 422
QY 422 KPDVNAFGTYILSARSSSLAPDSSFWANHDSKAYMGTSMATPIVAGNVAQLREHFVKR 481
DB 423 KPDVNAFGTYILSARSSSLAPDSSFWANHDSKAYMGTSMATPIVAGNVAQLREHFVKR 482
QY 482 GTPPKPSLLKALINAGADXXGLGYPNGNGMGWRTLLDKSLNVAAYNBSXSLSTSOKATYX 541
DB 483 GTPPKPSLLKALINAGADIGLGYPNGNGMGWRTLLDKSLNVAAYNBSXSLSTSOKATYX 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDTVITAPNGTYXGNDPFXKXXNMWGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDTVITAPNGTYXGNDPFXKXXNMWGR 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVPGPQXPSLAIVN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVPGPQXPSLAIVN 640

y.

RESULT 8
US-09-920-954-8
; Sequence 8, Application US/0920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-08-03
```

```
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PR
; ORGANISM: Bacillus sp.
US-09-920-954-8

Query Match
Best Local Similarity 98.8%; Score 3021; DB 3; Length 640;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFSLVLSAAAILSTVALKNPSAGXARFPLDFKGIQTTTDXKGFSGKXQOTGAAPL 61
DB 3 KKKKVFSLVLSAAAILSTVALSNPSAGARNPDLDFKGIQTTTDAKGFSGKXQOTGAAPL 62
QY 62 VSEENVKLGKGLKKLETPANNKLIHQPNPILBETKQLEXTGAKILDYIPDYAYIV 121
DB 63 VSEENVKLGKGLKKLETPANNKLIHQPNPILBETKQLEXTGAKILDYIPDYAYIV 122
QY 122 EYEGDVSKXXXXIEHVESVEPYLPYXIDPOLFTKGASLVKAXALDTQXNKEVOLRGI 181
DB 123 EYEGDVSKATSTIEHVESVEPYLPYRIDPOLFTKGASLVKAVALDTQXNKEVOLRGI 182
QY 182 EXIAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIVAADTGLDGT 241
DB 183 EQLAQPAISNDVLYITAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIVAADTGLDGT 242
QY 242 RNDSSWHEAFRGKITLVALGRTNNANDTNGHTHVASVLGNKGXTNKGMAPOANLVFQS 301
DB 243 RNDSSWHEAFRGKITLVALGRTNNANDTNGHTHVASVLGNKGXTNKGMAPOANLVFQS 302
QY 302 IMDSXGGLGGLPSNQTLPQXASAGARIHTNSWGAANVGAAYTTDSRNVDYVRKNDMTI 361
DB 303 IMDSXGGLGGLPSNQTLPQASAGARIHTNSWGAANVGAAYTTDSRNVDYVRKNDMTI 362
QY 362 LFAAGNEXPNGGTISAPGAKNAITVGATENLRPSGSYADNINHVAQSSSRGPTDGR 421
DB 363 LFAAGNEXPNGGTISAPGAKNAITVGATENLRPSGSYADNINHVAQSSSRGPTDGR 422
QY 422 KPDVNAFGTYILSARSSSLAPDSSFWANHDSKAYMGTSMATPIVAGNVAQLREHFVKR 481
DB 423 KPDVNAFGTYILSARSSSLAPDSSFWANHDSKAYMGTSMATPIVAGNVAQLREHFVKR 482
QY 482 GTPPKPSLLKALINAGADXXGLGYPNGNGMGWRTLLDKSLNVAAYNBSXSLSTSOKATYX 541
DB 483 GTPPKPSLLKALINAGADIGLGYPNGNGMGWRTLLDKSLNVAAYNBSXSLSTSOKATYX 542
QY 542 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDTVITAPNGTYXGNDPFXKXXNMWGR 601
DB 543 FTATAGKPLKISLVWSDAPASTTASVTLVNDLDTVITAPNGTYXGNDPFXKXXNMWGR 602
QY 602 NNVENVFINKPQSGTYTIEVOAYNVPGPQXPSLAIVN 639
DB 603 NNVENVFINKPQSGTYTIEVOAYNVPGPQXPSLAIVN 640

y.

RESULT 9
US-10-784-870-8
; Sequence 8, Application US/10784870
; Publication No. US2004042837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
```

APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIORITY APPLICATION NUMBER: US/09/509,814
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: PCT/JP98/04528
PRIORITY FILING DATE: 1998-10-07
PRIORITY APPLICATION NUMBER: JP 9-274570
PRIORITY FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-8

Query Match 98.8%; Score 3021; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 2,9e-260;
Matches 595; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

QY 2 RKKKVFSLVSAALISTVALXNPAGKARXEDLPKGIQTTDDXGKSPKQOTGAANFL 61
DB 3 KKKKVFSLVSAALISTVALXNPAGKARXEDLPKGIQTTDDXGKSPKQOTGAANFL 62
QY 62 VSEENVKLXKGLKLETPANNKLIHXQNGPILEETKQLEKXGAKLIDYIPYAYIV 121
DB 63 VSEENVKLXKGLKLETPANNKLIHXQNGPILEETKQLEKXGAKLIDYIPYAYIV 122
QY 122 EYEGDVXKXXXIHEVESVEPYLPYXIDPOLFTKGASXLVYAKALDTKXNKKEVOLRGI 181
DB 123 EYEGDVXKXXXIHEVESVEPYLPYXIDPOLFTKGASXLVYAKALDTKXNKKEVOLRGI 182
QY 182 EXIAQXXXNDXYITAKREYKVMNDVARGIVKADVAOSSVGLXQGGIVNADTGLDNG 241
DB 183 EYEGDVXKXXXIHEVESVEPYLPYXIDPOLFTKGASXLVYAKALDTKXNKKEVOLRGI 182
QY 242 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTAVAGSVLNGXGTXNKGMAPQANLVPQS 301
DB 243 RNDSSMHEAFRGKITALVALGRTNNANDTNGHGTAVAGSVLNGXGTXNKGMAPQANLVPQS 302
QY 302 IMDSYGGJGGLPSNQTLPFQKXSGARITHNSMGAANVGAAYTTDSRVNDYVRKNDWTI 361
DB 303 IMDSYGGJGGLPSNQTLPFQKXSGARITHNSMGAANVGAAYTTDSRVNDYVRKNDWTI 362
QY 362 LPAAGNEXPNGGTSAPGAKNAIVGATEMLRPSFGVADNINHVAFSSRGPTKDGRI 421
DB 363 LPAAGNEXPNGGTSAPGAKNAIVGATEMLRPSFGVADNINHVAFSSRGPTKDGRI 422
QY 422 KPDVVAPEGTXILISARSSILAPDSSFWANHDSKAYVNGGTSMATPIYAGVNAQLREHFVNQR 481
DB 423 KPDVVAPEGTXILISARSSILAPDSSFWANHDSKAYVNGGTSMATPIYAGVNAQLREHFVNQR 482
QY 482 GTPPEPSLLKALIGAADKGLGYNGNQGWGRVTLDSLVNAVYNESSXISTSQKATYX 541
DB 483 GTPPEPSLLKALIGAADKGLGYNGNQGWGRVTLDSLVNAVYNESSXISTSQKATYX 542
QY 542 FTATGKRLKTLVWSDAPASTTAVTLVNDLVTITAPNGXYVYGNPKPKXKXNMGR 601
DB 543 FTATGKRLKTLVWSDAPASTTAVTLVNDLVTITAPNGXYVYGNPKPKXKXNMGR 602
QY 602 NNVENVPINXPGSGTYTTEVQAYNPVPGPQXSLAIYN 639
DB 603 NNVENVPINXPGSGTYTTEVQAYNPVPGPQXSLAIYN 640

RESULT 10
US-09-920-954-1
Sequence 1, Application US/09920954

Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUMISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIORITY APPLICATION NUMBER: 09/509,814
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: PCT/JP98/04528
PRIORITY FILING DATE: 1998-10-07
PRIORITY APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURES:
NAME/KEY: misc feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid

```
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (596)..(596)
```

```
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-1
```

Query Match 98.6%; Score 3017; DB 3; Length 639;
Best Local Similarity 100.0%; Pred. No. 6,6e-260;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKKKVFLSVLSAAAILSTVALNPSAGAKRPDLDFKGIQTTHXKXGSKOXQTGAAP 60
DB 1 MKKKVFLSVLSAAAILSTVALNPSAGAKRPDLDFKGIQTTHXKXGSKOXQTGAAP 60
QY 61 LVSENVKLXKGLXKKLETPVANNKLIHQFNGLIETKOKLXTGAAILDYIPYAYI 120
DB 61 LVSENVKLXKGLXKKLETPVANNKLIHQFNGLIETKOKLXTGAAILDYIPYAYI 120
QY 121 VEYEGDVASXXXXIHVESVEPEYLPYXIDPOLFTKGASXLVKAXALDTKXNKVEQLNG 180
DB 121 VEYEGDVASXXXXIHVESVEPEYLPYXIDPOLFTKGASXLVKAXALDTKXNKVEQLNG 180
QY 181 IEXIAQXXXXNDVXYITAKPEYKVNNDVARGIVKADVAQSSYGLVGQGIYAVADTGLDT 240
DB 181 IEXIAQXXXXNDVXYITAKPEYKVNNDVARGIVKADVAQSSYGLVGQGIYAVADTGLDT 240
QY 241 GRNDSMHEAFRGKLTALYALGRTNNANDTNHGHVAGSVGNGKTKMGAPQANTVPO 300
DB 241 GRNDSMHEAFRGKLTALYALGRTNNANDTNHGHVAGSVGNGKTKMGAPQANTVPO 300
QY 301 SIMDSXGGLGGLPSNLQTLFSAQXAGARIHNSWGAANVGAYTTDSRVDDYRKNDMT 360
DB 301 SIMDSXGGLGGLPSNLQTLFSAQXAGARIHNSWGAANVGAYTTDSRVDDYRKNDMT 360
QY 361 ILFAAGNEXPNGGTSAPGTAGNAITVGATENLRPSFSYADNINHVAFSSRGPTKGR 420
DB 361 ILFAAGNEXPNGGTSAPGTAGNAITVGATENLRPSFSYADNINHVAFSSRGPTKGR 420
QY 421 IKPDVMAFGTILSRSSSLAPDSSFMANHDSRYATMGTSMAPIYVAGVADLRHPFVN 480
DB 421 IKPDVMAFGTILSRSSSLAPDSSFMANHDSRYATMGTSMAPIYVAGVADLRHPFVN 480
QY 481 RGITPKPSILKALJAGAADXGLGYPNGNGGRTVLDSLVAVYVNSSXLSSTOKATY 540
DB 481 RGITPKPSILKALJAGAADXGLGYPNGNGGRTVLDSLVAVYVNSSXLSSTOKATY 540
QY 541 XFTATAGRPKLKSLVMSDAPASTTASVTLVNDLIVITAPNGTXVVGNDPXXPXXNDMG 600
DB 541 XFTATAGRPKLKSLVMSDAPASTTASVTLVNDLIVITAPNGTXVVGNDPXXPXXNDMG 600
QY 601 RNNVENVPIINXPOSSTYITVQAVNVPGVQXFSLSAIYN 639
DB 601 RNNVENVPIINXPOSSTYITVQAVNVPGVQXFSLSAIYN 639
```

RESULT 11
US-10-784-870-1
Sequence 1, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUOOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGIYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT

```

CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PR1
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (126)..(126)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid

```


OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (103)..(103)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (106)..(106)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (129)..(129)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (131)..(131)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (132)..(132)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (133)..(133)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (134)..(134)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (147)..(147)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (149)..(149)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (161)..(161)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (166)..(166)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (173)..(173)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (184)..(184)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (188)..(188)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (189)..(189)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (190)..(190)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (195)..(195)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (287)..(287)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (307)..(307)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (325)..(325)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (370)..(370)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (432)..(432)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (502)..(502)	
OTHER INFORMATION: Xaa	is any amino acid
NAME/KEY: misc feature	
LOCATION: (532)..(532)	
OTHER INFORMATION: Xaa	is any amino acid

```

NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-2

```

```

Query Match      98.3%; Score 3007; DB 3; Length 640;
Best Local Similarity 100.0%; Pred. No. 5.2e-259;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	3	KKRVLSVLSAAALISTVALKNPSGAXRXPDLDPKGIQTTTDXKXGSKQXQXGAAPFLV	62
Db	4	KKRVLSVLSAAALISTVALKNPSGAXRXPDLDPKGIQTTTDXKXGSKQXQXGAAPFLV	63
Oy	63	ESENVVLXKGLKKKLETPANNKLIHXOPNPILEETKQLEKXGILDIYIPYAYIVE	1222
Db	64	ESENVVLXKGLKKKLETPANNKLIHXOPNPILEETKQLEKXGILDIYIPYAYIVE	123
Qy	123	YEGDVASXXXXLIEHVESVEPYLPYXIDIPOLFTKGASXLVAKALDTKQXNKEVOJRGIE	1822
Db	124	YEGDVASXXXXLIEHVESVEPYLPYXIDIPOLFTKGASXLVAKALDTKQXNKEVOJRGIE	183
Qy	183	XIAQXXXXNDYYIITRAKPEYKMDVAVGIKADVAOSSYGLYQGGIIVAVADTGLDTR	242
Db	184	XIAQXXXXNDYYIITRAKPEYKMDVAVGIKADVAOSSYGLYQGGIIVAVADTGLDTR	243
Qy	243	NDSMHEAFRGKITLVALGRTNNANDPNGHTVHAGSVLGNKXTNGMAPOANLVPQSI	302
Db	244	NDSMHEAFRGKITLVALGRTNNANDPNGHTVHAGSVLGNKXTNGMAPOANLVPQSI	303
Qy	303	MDXGGLGGLPSNLQTLPSQAXSAGARLHTNSWGAANVGAATYDTSRVVDYVRKNDWTIL	362
Db	304	MDXGGLGGLPSNLQTLPSQAXSAGARLHTNSWGAANVGAATYDTSRVVDYVRKNDWTIL	363
Qy	363	PAAGNXPNGGRTISAPGRKAKALITYGATEENRPSFGSTADNINNVAPSSRGPIKGRIRK	422
Db	364	PAAGNXPNGGRTISAPGRKAKALITYGATEENRPSFGSTADNINNVAPSSRGPIKGRIRK	423
Qy	423	PDVMAEGXTIILSARSSLAPDSSFWANHDHCKTAYVNGGTSMTAPIYAGVAVOLREHFVKNRG	482
Db	424	PDVMAEGXTIILSARSSLAPDSSFWANHDHCKTAYVNGGTSMTAPIYAGVAVOLREHFVKNRG	483
Qy	483	ITPKPSILKALAIAGADGXGLGYPNGQNGWGRVTLIDJLSLVAYVNESSXLSTSQATYXF	542
Db	484	ITPKPSILKALAIAGADGXGLGYPNGQNGWGRVTLIDJLSLVAYVNESSXLSTSQATYXF	543
Qy	543	TATAGPPLKISLVMSDAPASTTASTYLVNDLVIITAPNGXTYYGNDPXXPKXXMNDGRN	602
Db	544	TATAGPPLKISLVMSDAPASTTASTYLVNDLVIITAPNGXTYYGNDPXXPKXXMNDGRN	603

QY 603 NVENVFNKPGSTYTIETVQAVNVPYGPQXFSIAIYN 639
DB 604 NVENVFNKPGSTYTIETVQAVNVPYGPQXFSIAIYN 640

RESULT 13
US-10-784-870-2
Sequence 2, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid

OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-2

```

Query Match 98.3%; Score 3007; DB 4; Length 640;
 Best Local Similarity 100.0%; Pred. No. 5.2e-259; Indels 0; Gaps 0;
 Matches 637; Conservative 0; Mismatches 0;

```

QY 3 KKKVSLVLSAAIISTVALXNPSAGXARXFDLDFKGIQTTTDXGFGSKQQTGAALV 62
DB 4 KKKVSLVLSAAIISTVALXNPSAGXARXFDLDFKGIQTTTDXGFGSKQQTGAALV 63
QY 63 ESENKXKXGKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 122
DB 64 ESENKXKXGKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 123

```

```

QY 123 YEGDVXSKXXIXIEHVESVEPILPYXXIDPOLPTKGASLYVQXALDPTKQXKXEQVLRGIE 182
DB 124 YEGDVXSKXXIXIEHVESVEPILPYXXIDPOLPTKGASLYVQXALDPTKQXKXEQVLRGIE 183
QY 183 XIAOXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGGQGIIVAADTGLDTR 242
DB 184 XIAOXXSNDVXYITAKPEYKVMNDVARGIVKADVAOSSYGLYGGQGIIVAADTGLDTR 243
QY 243 NDSMHEAFRGKITVALGRTNANDTNGHGTIVAGSVLNGXTNKMAPQANLVFQSI 302
DB 244 NDSMHEAFRGKITVALGRTNANDTNGHGTIVAGSVLNGXTNKMAPQANLVFQSI 303
QY 303 MDSXGGLGGLPSNIQTLFSAQXASAGARHTMSKGAUVNGATTPTSRNDVYRKNDMTIL 362
DB 304 MDSXGGLGGLPSNIQTLFSAQXASAGARHTMSKGAUVNGATTPTSRNDVYRKNDMTIL 363
QY 363 FAAGNEXPNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQPSRGPYKGRIR 422
DB 364 FAAGNEXPNGGTISAPGTAKNAITVGATENTLRPSFGSYADNINHVAQPSRGPYKGRIR 423
QY 423 PDVAPGTXTLSARSSILAPDSSFWANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKRG 482
DB 424 PDVAPGTXTLSARSSILAPDSSFWANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKRG 483
QY 483 ITPKPSILKALILGAADXGYPNGNGKWRVTLDSLSNVAAYNESXISTSQATYXF 542
DB 484 ITPKPSILKALILGAADXGYPNGNGKWRVTLDSLSNVAAYNESXISTSQATYXF 543
QY 543 TATAGKPLKISLVNSDAPASTTASVTLVNDLDTLITAPNGTXYYGNDPFXKXXKXMDGRN 602
DB 544 TATAGKPLKISLVNSDAPASTTASVTLVNDLDTLITAPNGTXYYGNDPFXKXXKXMDGRN 603
QY 603 NVENVFINXPOSGTYTIEVQAYNVPGPQXFSIAVN 639
DB 604 NVENVFINXPOSGTYTIEVQAYNVPGPQXFSIAVN 640

```

RESULT 14
 US-09-985-689A-1
 ; Sequence 1, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAMA, AKINORI
 ; APPLICANT: KAGIYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAKETI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483USO
 ; CURRENT APPLICATION NUMBER: US/09/985, 689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 434
 ; TYPE: PRT
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 3; Length 434;
 Best Local Similarity 96.3%; Pred. No. 2.9e-183; Indels 0; Gaps 0;
 Matches 418; Conservative 0; Mismatches 16;

```

QY 206 NDVARGIVKADVAOSSYGLYGGQGIIVAADTGLDTRNDSMHEAFRGKITVALGRTN 265

```

Db 1 NDVARGIVKADVAQSSYGLYGCGQIVAVADTGLDTGRNDSMHBAFRGKITLALYALGRTN 60
Qy 266 NADNTGCHGVAGSVLGNKXTTKGMAPOANLVPQSIMSXGGLGGLPSNLQTLFEOAXS 325
Db 61 NADNTGCHGVAGSVLGNKXTTKGMAPOANLVPQSIMSXGGLGGLPSNLQTLFEOAXS 120
Qy 326 AGARHTNSWGAANVGAATTTDSRNVDYVRKNDMTLLFAAGNEXPGNGTISAGTAKNAI 385
Db 121 AGARHTNSWGAANVGAATTTDSRNVDYVRKNDMTLLFAAGNEXPGNGTISAGTAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAFGTYILSARSSLPADSSP 445
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAFGTYILSARSSLPADSSP 240
Qy 446 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALILAGAADYGLGY 505
Db 241 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALILAGAADYGLGY 300
Qy 506 PNGNQGMRVTLDKSLNVAAYVNESSXLSTSQKATYXFTATAGKPKLSLWSDAPASTTA 565
Db 301 PNGNQGMRVTLDKSLNVAAYVNESSXLSTSQKATYXFTATAGKPKLSLWSDAPASTTA 360
Qy 566 SVTLVNDLVLVTAPNGTYVGNDFPKXPKXNMDDGRNNVENVFINXPOSGTYTIEVOAYN 625
Db 361 SVTLVNDLVLVTAPNGTYVGNDFPKXPKXNMDDGRNNVENVFINXPOSGTYTIEVOAYN 420
Qy 626 VEVGPQXFSLATVN 639
Db 421 VEVGPQXFSLATVN 434

RESULT 15

US-09-985-689A-2
; Sequence 2, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. No. 2.9e-183;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 206 NDVARGIVKADVAQSSYGLYGCGQIVAVADTGLDTGRNDSMHBAFRGKITLALYALGRTN 265
Db 1 NDVARGIVKADVAQSSYGLYGCGQIVAVADTGLDTGRNDSMHBAFRGKITLALYALGRTN 60
Qy 266 NADNTGCHGVAGSVLGNKXTTKGMAPOANLVPQSIMSXGGLGGLPSNLQTLFEOAXS 325
Db 61 NADNTGCHGVAGSVLGNKXTTKGMAPOANLVPQSIMSXGGLGGLPSNLQTLFEOAXS 120
Qy 326 AGARHTNSWGAANVGAATTTDSRNVDYVRKNDMTLLFAAGNEXPGNGTISAGTAKNAI 385

Db 121 AGARHTNSWGAANVGAATTTDSRNVDYVRKNDMTLLFAAGNEXPGNGTISAGTAKNAI 180
Qy 386 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAFGTYILSARSSLPADSSP 445
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAFGTYILSARSSLPADSSP 240
Qy 446 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALILAGAADYGLGY 505
Db 241 WANHDSKYAYMGTSMAATPIVAGNVAQLREHFVKRNGITPKPSLLKALILAGAADYGLGY 300
Qy 506 PNGNQGMRVTLDKSLNVAAYVNESSXLSTSQKATYXFTATAGKPKLSLWSDAPASTTA 565
Db 301 PNGNQGMRVTLDKSLNVAAYVNESSXLSTSQKATYXFTATAGKPKLSLWSDAPASTTA 360
Qy 566 SVTLVNDLVLVTAPNGTYVGNDFPKXPKXNMDDGRNNVENVFINXPOSGTYTIEVOAYN 625
Db 361 SVTLVNDLVLVTAPNGTYVGNDFPKXPKXNMDDGRNNVENVFINXPOSGTYTIEVOAYN 420
Qy 626 VEVGPQXFSLATVN 639
Db 421 VEVGPQXFSLATVN 434

Search completed: April 7, 2006, 18:03:40
Job time : 120.407 secs

This Page Blank (uspto)

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:59:52 ; Search time 16.4871 Seconds

(without alignments)
1208.926 Million cell updates/sec

Title: US-10-784-870-1

Perfect score: 3059

Sequence: 1 MRKKVPLSVLSAAALSTV.....EYQAVNPVGPQXFLATVN 639

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100*

Listing first 45 summaries

Database : Published Applications AA New:*
1: /SIDS5/prodata/1/pubppa/US08_NEW_PUB pep:*
2: /SIDS5/prodata/1/pubppa/US06_NEW_PUB pep:*
3: /SIDS5/prodata/1/pubppa/US07_NEW_PUB pep:*
4: /SIDS5/prodata/1/pubppa/US09_NEW_PUB pep:*
5: /SIDS5/prodata/1/pubppa/US10_NEW_PUB pep:*
6: /SIDS5/prodata/1/pubppa/US11_NEW_PUB pep:*
7: /SIDS5/prodata/1/pubppa/US12_NEW_PUB pep:*
8: /SIDS5/prodata/1/pubppa/US13_NEW_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	276.5	9.0	802	US-10-510-386-2	Sequence 2, Appli
2	274.5	9.0	874	US-10-510-386-28	Sequence 28, Appli
3	274.5	9.0	1047	US-10-510-386-200	Sequence 200, Appl
4	249.5	8.2	379	US-11-156-062-23	Sequence 23, Appl
5	242	7.9	272	US-11-020-602-236	Sequence 236, Appl
6	239	7.8	269	US-11-020-602-6	Sequence 6, Appli
7	237	7.7	382	US-11-020-602-2	Sequence 2, Appli
8	227.5	7.4	274	US-11-156-062-14	Sequence 14, Appli
9	227.5	7.4	275	US-11-065-943-54	Sequence 54, Appli
10	227.5	7.4	275	US-11-020-602-3	Sequence 3, Appli
11	226.5	7.4	274	US-11-156-062-12	Sequence 12, Appli
12	225.5	7.4	274	US-11-020-602-5	Sequence 5, Appli
13	225.5	7.3	274	US-11-156-062-4	Sequence 4, Appli
14	221.5	7.2	274	US-11-156-062-10	Sequence 10, Appli
15	220.5	7.2	274	US-11-156-062-6	Sequence 6, Appli
16	220.5	7.2	275	US-11-156-062-8	Sequence 8, Appli
17	217.5	7.1	274	US-11-020-602-4	Sequence 4, Appli
18	217.5	7.1	274	US-11-156-062-16	Sequence 16, Appli
19	217.5	7.1	274	US-11-156-062-18	Sequence 18, Appli
20	216.5	7.1	274	US-11-156-062-2	Sequence 2, Appli
21	210.5	6.9	1052	US-11-020-602-208	Sequence 208, Appl
22	204.5	6.7	1432	US-10-510-386-218	Sequence 218, Appl
23	198.5	6.5	280	US-11-020-602-209	Sequence 209, Appl
24	197.5	6.5	1647	US-11-052-554A-260	Sequence 260, Appl
25	179	5.9	740	US-11-096-568A-24714	Sequence 24714, A

26	179	5.9	777	US-11-096-568A-24713	Sequence 24713, A
27	179	5.9	790	US-11-096-568A-24712	Sequence 24712, A
28	166.5	5.4	591	US-10-510-386-22	Sequence 22, Appli
29	160	5.2	722	US-11-096-568A-31863	Sequence 31863, A
30	160	5.2	757	US-11-096-568A-31862	Sequence 31862, A
31	160	5.2	798	US-11-096-568A-31861	Sequence 31861, A
32	152	5.0	733	US-11-096-568A-24028	Sequence 24028, A
33	152	5.0	759	US-11-096-568A-24027	Sequence 24027, A
34	152	5.0	764	US-11-096-568A-24026	Sequence 24026, A
35	151	4.9	659	US-11-096-568A-17895	Sequence 17895, A
36	151	4.9	791	US-11-096-568A-17894	Sequence 17894, A
37	151	4.9	791	US-11-096-568A-31251	Sequence 31251, A
38	150	4.9	791	US-11-218-986-2	Sequence 2, Appli
39	148.5	4.9	794	US-10-821-234-1176	Sequence 1176, Ap
40	148.5	4.9	820	US-11-096-568A-23983	Sequence 23983, A
41	143	4.7	680	US-11-096-568A-23982	Sequence 23982, A
42	143	4.7	470	US-10-873-528-3	Sequence 3, Appli
43	134	4.4	617	US-11-096-568A-31252	Sequence 31252, A
44	132.5	4.3	617	US-11-096-568A-23984	Sequence 23984, A
45	130	4.2	418	US-11-096-568A-23984	Sequence 23984, A

ALIGNMENTS

RESULT 1									
US-10-510-386-2									
; Sequence 2, Application US/10510386									
; Publication No. US20050244922A1									
; GENERAL INFORMATION:									
; APPLICANT: Andersen, Jens Tonne									
; APPLICANT: Clausen, Ib Groth									
; APPLICANT: Jorgensen, Steen Troels									
; APPLICANT: Olsen, Peter Bjarke									
; APPLICANT: Rasmussen, Michael Dolberg									
; TITLE OF INVENTION: Improved Bacillus Host Cell									
; FILE REFERENCE: 10294.204-US									
; CURRENT APPLICATION NUMBER: US/10/510.386									
; CURRENT FILING DATE: 2004-10-04									
; NUMBER OF SEQ ID NOS: 248									
; SOFTWARE: Patent in version 3.3									
; SEQ ID NO 2									
; LENGTH: 802									
; TYPE: PRT									
; ORGANISM: Bacillus licheniformis									
US-10-510-386-2									
Query Match									
Beet Local Similarity 9.0%; Score 276.5; DB 6; Length 802;									
Matches 160; Conservative 82; Mismatches 226; Indels 265; Gaps 31;									
QY	126	DVKSXXXXIHEVSEVPEYLPYXXIDPOLFTGASXLVAXALDT---	KOXNREVOALRGIE	182					
DB	49	DVTSQKQTTVVELKESLAKADGKQTKASLTKAKAKAREVD-RVPS	107						
QY	183	XIAQXXNSNDVXYITAKPEYVM-----	NDVARGIVRAD	216					
DB	108	GFSMGLPASEIRPKLAVKEKAVVFNATYKPDVSGKQVTLAADAIVPQMKASAPFIGAD	167						
QY	217	VAQSSVGIYGGQIVAVADTGLDTCGRND--SSMHEAFRKK--ITLALYALGRTNNADTNG-	272						
DB	168	QAWNS-GVTGKGIKAVAVDVTGDTYTHPLKKNFGPKYDVPDNDYDDEFTPTGPRGGA	226						
QY	273	--HGTHTVAGSVLGGNGXTKNGAPQNLVFOGIMDSXGAGLPSNLOTLPQASAGAGARI	330						
DB	227	TDHGHVAGTIAANGQI-KGVAPRATILAYVLGRG--SGTTEVVIAGIEKAVADGAKV	283						
QY	331	HTNSKGAAVNGA-YTTDSRNVDDYVRKNDITILPAGNEXXNGGTTISAPGAKAIVTGA	389						
DB	284	MNLSTGKSLNSPDVATSI--ALDWAMEGVAVLVTSNGSGENMTVSGPSTSRVAISVGA	341						
QY	390	TE-----NLRPSFGS-----	399						

```

Db 342 SQLPYNEXSVTLPSYSAKWVGQEERKOLEALNGQVELVEAGLQADDPFGSKDYKGYA 401
Qy 400 -----YADNINH----- 406
Db 402 VIORGVIPFVDKAEBAKAGAIYINNATGEIBANWGMVPTVKLSKEGEKLVQOI 461
Qy 407 -----VAQSSRGPTKO-GRIKPDVMAFGTXIISARSLSLPDSSF 445
Db 462 KEKGKSVVSFFLDKCKLGETTISFSSRGFVMDTMMIKPDVSAFGVNIYSTIPT----- 514
Qy 446 WANHDSK-----YAWGTSMTATPIVAGNVAQUREHFVKNRGITP--KPSLLKALIAAGNA 499
Db 515 ---HDPKAPYVYGSQGTSMASPHVAGTAALIKQ-----AKPDMTEQIGVLMNTAB 564
Qy 500 ---DXGIGYFNGNGMGRVTLDSKLVN-----AYNESS 530
Db 565 KLTDBNGPFLPHNTQAGSIRIRLMBALKASSIYTPGSHSYGFLKDKGQKQKKAFTIEN- 623
Qy 531 XLSTSQKA---TYXFTATAGKPLKISLWSD---APASTT---ASVTLVNDLIVITAPN 581
Db 624 -LSHRKAIYQLEYSGKGTG---ITVSGTERVVPVPAQGTGAALKVTVNSAKTYAGTYE 677
Qy 582 GTXYVGNDFKXXPKXXNMGRNVE---NVFINKP-----QSGTYTIEVOAY 624
Db 678 GTVYIRB-----DQRKVAEIPLLIIVKEPDYPRVTSVTVPEPAKQAGATTB--AY 725
Qy 625 NVFVGPQKPSLAI 637
Db 726 -LPGAEELAFV 737

RESULT 2
US-10-510-386-28
; Sequence 28, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Oleen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 28
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-28

```

```

Query Match          9.0%; Score 274.5; DB 6; Length 874;
Best Local Similarity 23.7%; Pred. No. 1.4e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;

```

```

Qy 13 AAAILSTVALXNPSAGAKRFDLPFGKIQTDDXXGFSKQQTGAAP-----LV 62
Db 213 AAPFISAKMLFNKSMRDSVYKDL---VQLKPLFADVAKNGQVSAVITNDQKASISLY 268
Qy 63 ESENVKXKXGKXKLETPANNKLIHQPNPILBETKQKLEKXG-AKILDIYIPDAYIV 121
Db 269 ETAPASVEPLKQQLDQVAKD--IGIEQLTG--SKVASVLEKAGMTASSAPENRIV 323
Qy 122 EYE-----GDVYKXXXXIEHVSVEP-----YLPYXXID-POLFTGASXLVKAHALD 168
Db 324 KXGKXKPSGFSKSAQS--SGVQALEPLGKSKTAFKMTVVENKESRSSGFRAAKQYQAA 382
Qy 169 TKQXKKEVQLRGIEIXIAQ-XXXSNDVXYITAKPEYKVNADVARGIVADVAQSSGYLYGQ 227
Db 383 ASKIAKMEVEFEVQVQYEALESRTQY---PYQWSLKNNGKNRAANADIQFEOQLKLMK 439

```

```

Qy 228 GQ-----IVAVADTGLDGRNDSMHEAFRGI---TALVALGRTNNAANDNGHGTIVAG 279
Db 440 GKXKLDVTIYAVVDGVGVDHTLADLS-----GSVKKDEGYNVGRTRADMMDNGHGTIVSG 493
Qy 280 SVLG---NGXTNKMAAPQANLVFQSIIMDSXGGLGSLPSNLQTLPSQAXSAGARIHTNSWG 336
Db 494 IIAAODHMFMAGINAVAKILPVKVLDSG--SGDPEQIANGIIVAADHAKYINLSLG 551
Qy 337 AAVNGAYTDSRNVDDYRKNDMTTILPAAGNEKPNNGTISAPGAKNAITGATENTRPS 396
Db 552 ---GPSRWEYALKYASKQVITIVATGNDGVS--EISYPASSKYLISGATNNL--- 602
Qy 397 FGSAVDNINHAQPSNRGPTDGRIKPDVMAFGTXIISARSLSLPDSSFMANHDSKAYM 456
Db 603 -----DLVSDYSNIGKGL-----DMVAPGTI---PSLVPDGN-----VTYM 636
Qy 457 GGTSMATPIVAGNVAQUREHFVKNRGITPKEPSLLKALIAAGAXGL---GYPNGN----- 509
Db 637 SGTSMAPHVAAAAGL-----LSQNPISLRFKQJASLITETTLADVAPEBQDNPRPDYDLD 691
Qy 510 -----QGMGRVTLDSKLVNAYNESSXLSSTSQKATYXFTATAGKPLKI 552
Db 692 IEPAAQIFGYDVFVSGWGRLVNFHAASVFEIMKQHPVILNRTAVTGTAKSGVTVKI 747

```

```

RESULT 3
US-10-510-386-200
; Sequence 200, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Oleen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 200
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-200

```

```

Query Match          9.0%; Score 274.5; DB 6; Length 1047;
Best Local Similarity 23.7%; Pred. No. 1.7e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;

```

```

Qy 13 AAAILSTVALXNPSAGAKRFDLPFGKIQTDDXXGFSKQQTGAAP-----LV 62
Db 213 AAPFISAKMLFNKSMRDSVYKDL---VQLKPLFADVAKNGQVSAVITNDQKASISLY 268
Qy 63 ESENVKXKXGKXKLETPANNKLIHQPNPILBETKQKLEKXG-AKILDIYIPDAYIV 121
Db 269 ETAPASVEPLKQQLDQVAKD--IGIEQLTG--SKVASVLEKAGMTASSAPENRIV 323
Qy 122 EYE-----GDVYKXXXXIEHVSVEP-----YLPYXXID-POLFTGASXLVKAHALD 168
Db 324 KXGKXKPSGFSKSAQS--SGVQALEPLGKSKTAFKMTVVENKESRSSGFRAAKQYQAA 382
Qy 169 TKQXKKEVQLRGIEIXIAQ-XXXSNDVXYITAKPEYKVNADVARGIVADVAQSSGYLYGQ 227
Db 383 ASKIAKMEVEFEVQVQYEALESRTQY---PYQWSLKNNGKNRAANADIQFEOQLKLMK 439
Qy 228 GQ-----IVAVADTGLDGRNDSMHEAFRGI---TALVALGRTNNAANDNGHGTIVAG 279
Db 440 GKXKLDVTIYAVVDGVGVDHTLADLS-----GSVKKDEGYNVGRTRADMMDNGHGTIVSG 493
Qy 280 SVLG---NGXTNKMAAPQANLVFQSIIMDSXGGLGSLPSNLQTLPSQAXSAGARIHTNSWG 336

```

Db 494 IIAAADNHFMSMAGINAAKILEPVKYLDSGG--SGDTEQIANGIIYAADHGAKVINTLSIG 551
Qy 337 AAVNGAYTTDSRNVDDYKAKNDMTILFAAGNEXPNCGTISAPCTAGATTVGATEMLRPS 396
Db 552 ---GYSYVMEYALFYAASKNVTIYAATGNDGVS--EASTPASSKYTILSVGATNNL--- 602
Qy 397 FCSYADINNHVAOPFSRSGFTKQGRIKPDVMAFGTXILSARSSLAPOSSFWANHDSKAYM 456
Db 603 ---DLYSDYNYGKGL-----DVAAPGTDL-----PSLVDPGN-----VTTM 636
Qy 457 GGTSMATPIVAGNVAQIAREHFVNKNGITPSPILKALJAGADKGL---GYPNGN----- 509
Db 637 SGTSMAPPHVAAAAGLL-----LSQNPSPKPKQIASLITETPADVAFBEQDNPNDPYDD 691
Qy 510 -----OCGRVTLDSKLSLVAAVYNSSKSLSTQKATYXTATAGKPLKI 552
Db 692 IEPAAQIPGYDFVSGWGLNLFVFAASVFELNMKVHPVLRHRAVGTAKSGVTVKI 747

RESULT 4

US-11-156-062-23
; Sequence 23, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wileland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueßer, Inken
; APPLICANT: Stenr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 379
; TYPE: PRY
; ORGANISM: Bacillus licheniformis
US-11-156-062-23

Query Match 8.2%; Score 249.5; DB 7; Length 379;
Best Local Similarity 23.7%; Pred. No. 4.1e-12;
Matches 117; Conservative 58; Mismatches 164; Indels 155; Gaps 18;

Qy 1 MRKKVPLSVLSAAAILSTVALXNPSAGK--ARXFDLPKGIQTTDDXXGFSKQKQTGA 58
Db 2 MRKKSFWLGMILTAFLVFTMAFSDSASAAQPAKNEKDY-----IYFSGVATAS- 52
Qy 59 AFLVESENVKYLKXGLKLETPANNKLIHQFNGPILEETKQXLEXTAKLIDYIPDA 118
Db 53 ---VKDIIKESGGKVDK-----QFR--IIAAAKAKDKKALKKVKNDPDVA 94
Qy 119 YIVVEGVDVXXXXXIEHVESVEPYLPXXYIDPOLFTKGSXLYVAKXALDTQXNKEVQL 178
Db 95 YVER-1-----DHVAHALAQTVPGIP----- 114
Qy 179 RGIEKIAQXXXSNDVXXYITAKPEYKAMDVAGIYKADVAQSSYGLYGQGOIVANADTGL 238
Db 115 -----LTKADRVQAO--GFKGANVVAVALDTGI 140
Qy 239 DTGRNDSMHEBAFRKITLVALGRTNNANDTNGHGTIVAGSV--LGNGXTNKGMAPQAN 296
Db 141 QASHDPLANVGG-----ASPIVAGKAYN--TDGNGHGTIVAGTVAALDNTTGVLGVAPSVS 193
Qy 297 LVFQSIIMDSXGGLGILPSNLQTLFSGQASAGARIHTNSGAAVNGAYTTDSRNVDDYK 356
Db 194 IYAVAVKLNSSG--SGSYSGIVSGIEWATNGMDVINMISLGA---SGSTAMKQAVDNVA 248

Qy 357 NDMTILFAAGNEXPNCG--GTISAPGTAKNAITVGATENLRPSFGSYADINNHVAOPSSRG 414
Db 249 KGVVVAALAGNCGSGNTNTIGYPAKTDVIAVGA-----VDSNSRPAFSVVG 297
Qy 415 PYKDGRIKPDVMAFGTXILSARSSLAPOSSFWANHDSKAYMGTSMATPIVAGNVA--- 471
Db 298 -----ALEFWAPAGGVSTYPT-----NTYATLNGTSMASPHVAAAALL 339
Qy 472 ---QUREHFVZOR 481
Db 340 SKHPNLSAQVNRK 353

RESULT 5

US-11-020-602-236
; Sequence 236, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 236
; LENGTH: 272
; TYPE: PRY
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hybrid of
; OTHER INFORMATION: Bacillus lencus and Bacillus amyloliquefaciens
US-11-020-602-236

Query Match 7.9%; Score 242; DB 7; Length 272;
Best Local Similarity 31.9%; Pred. No. 1e-11;
Matches 84; Conservative 30; Mismatches 101; Indels 48; Gaps 10;

Qy 213 VQADVAQSSYGLYGQGOIVANADTGLDTGRNDSMHEBAFRKITLVALGRTNNANDTNG 272
Db 11 VQAPAAHNR--GLTSGGVVAVLDTGIST-----HPDLNIRGASFPVGBP--STQDNG 61
Qy 273 HGTIVAGSV--LGNGXTNKGMAPQANLVFQSIIMDSXGGLGILPSNLQTLFSGQASAGARI 330
Db 62 HGTIVAGTIALMNNISIGLVGVAPSAELVAVKVLGASG--SGSVSSIAQGLFWAGNNGMY 119
Qy 331 HTNSGAAVNGAYTTDSRNVDDYKAKNDMTILFAAGNEXPNCG--GTISAPGTAKNAITVG 388
Db 120 IMMSLGGSGAAL-----KAAVDKAVASGVVVAALAGNCTSSSSTVIGPKYPSVIAVG 175
Qy 389 ATENLRPSFGSYADINNHVAOPFSRSGFTKQGRIKPDVMAFGTXILSARSSLAPOSSFWAN 448
Db 176 A-----VDSNQRASFSVGP-----BLDVNAPGVISIGTLPG----- 208
Qy 449 HDSKTAVMGTSMATPIVAGNVA 471
Db 209 --NKYGAVNGTSMASPHVAGAA 229

RESULT 6

US-11-020-602-6
; Sequence 6, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND

```

; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GC527C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Bacillus lentus
US-11-020-602-6

Query Match      7.8%; Score 239; DB 7; Length 269;
Best Local Similarity 30.7%; Pred. No. 1.7e-11;
Matches 90; Conservative 30; Mismatches 91; Indels 82; Gaps 14;

QY 213 VKADVAAOSSYGLYGGGQIVAAVDGLDTCGRNDSMHEAFRGKITLALYALGRTNMNDPNG 272
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 11 VQAPPAHNR-GLTSGGVKVAALDTGIST-----HPDLNIRGASFPVGPB-STQDNG 61
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 273 HGTIVAGSV--LGNGXTNKGMAPQANLVFQSIIMDSXGGLPSNLTFLSQAKSAGARI 330
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 62 HGTIVAGTIALNNSIGVLGVAPELVAVKVLGASG--SGSVSSIAQGLEWAGNNGHHV 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 331 HTNSWGA-----AVNGATTDSSRVVDYVRKNDITLFAAGNEKXPNGGTISAPGTA 381
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 120 ANLSIGSPSPATLEQAVMSA---TSRGV-----LVVAASGNS--GAGSISYPPARY 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 382 KNAIVGATE--NLSPFGSYADNINHVAFSSRGPTKGRKRPVMAPGTIIISARSSL 439
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 166 ANAAVAVGATDQNNNPAFSFGYAGI-----DIVAPGVNVOSTYRG- 205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 440 APDSFMANHDSKYAMGTSMTPTIVAGNVA-----QLRHFVFN 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 206 -----GTVASLNGTSMATPHVAGAAALVYQKQNPMSNGVDIRNH-LKN 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-11-020-602-2
; Sequence 2, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Estcell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GC527C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (163)...(163)
; OTHER INFORMATION: Xaa = Asn or Pro
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (164)...(164)
; OTHER INFORMATION: Xaa = Pro or Asn
; FEATURE:
; NAME/KEY: VARIANT
```

```

; LOCATION: (168)...(168)
; OTHER INFORMATION: Xaa = Asn or Asp
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (195)...(195)
; OTHER INFORMATION: Xaa = Ala or Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (196)...(196)
; OTHER INFORMATION: Xaa = Ser or Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (205)...(205)
; OTHER INFORMATION: Xaa = Ala or Asp
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (206)...(206)
; OTHER INFORMATION: Xaa = Asp or Ala
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (265)...(265)
; OTHER INFORMATION: Xaa = Thr or Ser
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (266)...(266)
; OTHER INFORMATION: Xaa = Ser or Thr
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (358)...(358)
; OTHER INFORMATION: Xaa = Gln or Glu
US-11-020-602-2

Query Match      7.7%; Score 237; DB 7; Length 382;
Best Local Similarity 23.8%; Pred. No. 4e-11;
Matches 119; Conservative 40; Mismatches 152; Indels 188; Gaps 17;

QY 1 MRKKVEFLSYLSAAAILSTVALXNPSAGXARXFDLPKGIQTITDXXGFSKQOTGAAP 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MRGKQWISLIFALALFTVAFGSTSSAQAG-----KSNQEKRYIGFQOTMSTMAA- 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 LVSEENVKXKKGKLETPANNKLIHQFNGPILEBYOXLEXTAKILDIYIPDYAYI 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 55 -----KKDDVISEKGVQKQFK--YVDAAGATLNEKAVKEIKKQPSVAYV 98
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 VEYEGDVAXXXXXIEHY-----ESVEPYLPXYXIDPQLFTFKGASXLYKAYALDTKQNK 175
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 99 EE-----DHVAHVAQSV-PY-----GVSQ-IKAPAL----- 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 VOLRGIEIXAQXXXSNDVXYITAKPEYKWMNDVARGIKADVAGSSYGLYGQGIIVADV 235
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 124 -----HSGGYTGSNVKVAVID 139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 236 TGLDTGRNDSMHEAFRGKITLALYALGRTNMNDNGHTIVAGSV--LGNGXTNKGMAP 293
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 140 SGID-----SSHDLKVAAGASWVPSEITXXFQDXNSHGTIVAGTVALNNSIGVLGVA 193
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 294 QANLVFQSIIMDSXGGLGGLPSNLTFLSQAKSAGARIHTMSWGAANVAGATTDSRND-- 351
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 194 SXKLYAVKVLGKXG-----SQQYSWINGIEIMALNNDVVI 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 352 -----DYVRKNDMTILFAAGNEKPNG--GTISAPGTAKAALTVGATEN 392
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 230 NMSLGGPFGSAAALKAADVXAVASGVVVVAAAAGNGXGSSSTVGVPKRYGSVIAVGA--- 286
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 393 LRPSFGSYADNINHVAFSSRGPTKGRIRKPDVMAPGTXILSARSSLAIPSSFMANHDSK 452
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 287 -----VDSNQRAFSFSGVP-----ELDVAPGVSVISQSTLPG-----NK 320
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 453 YAVMGTSMATPTIVAGNVA 471
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 321 YGATNGTSMASPHVAGAA 339
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 8
US-11-156-062-14
; Sequence 14, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Mieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueger, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-14
```

```
Query Match          7.4%; Score 227.5; DB 7; Length 274;
Best Local Similarity 30.2%; Pred. No. 1.4e-10;
Matches 85; Conservative 35; Mismatches 108; Indels 53; Gaps 11;
```

```
QY 212 YKADVAQSSYGLYGGQCIYAAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 271
DB 10 LKAPKVOAQ-GFKGNVVAALDTGIQASHPLNVVG-----ASFVAGEAYNA-DDN 61
QY 272 GHTHVAQSV--LGNCKTNKMAPOANLVFQSIMDSXGGLGSPNLQTLFSGQAXSAGAR 329
DB 62 GHTHVAQSV--LGNCKTNKMAPOANLVFQSIMDSXGGLGSPNLQTLFSGQAXSAGAR 329
QY 330 HTNSGAAVNGAYTTDSRNVDDYRKNDMTLLFAAGNEXPN--GTISAPGTAKNAITV 387
DB 120 VINMSIGA--SGSTAMQAVDNAAVAGVVAAGSSGNTTIGVPAKYDSVI 176
QY 388 GATENLRPSFGSYADNINHAQPSRGPPTKDGRIKPDVAPGTIILSARSSLAPDSSFWA 447
DB 177 GA-----VDSNSRASFSSVG-----ABLEWAPGAGVYSTPT----- 210
QY 448 HNDKXYVMGSTMATPIYAGNVA-----QLRHFVKNR 481
DB 211 ---NTYATWNGTSMASPHVAGAAALILSKHPNLASQVNR 248
```

```
RESULT 9
US-11-065-943-54
; Sequence 54, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:
; APPLICANT: JESTIN, JEAN-LUC
; APPLICANT: VICHIER-GUERRE, SOPHIE
; APPLICANT: FERRIS, STEPHANE
; TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
; TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES,
; FILE REFERENCE: 266426USOXCI
; CURRENT APPLICATION NUMBER: US/11/065,943
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/787,219
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 54
; LENGTH: 275
```

```
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-065-943-54
```

```
Query Match          7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;
```

```
QY 213 YKADVAQSSYGLYGGQCIYAAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 272
DB 11 IKAP-ALHSQYTGSNVVAVIDSGID-----SSHPLKVAAGSAMVPSRTNPFQDNNS 63
QY 273 GHTHVAQSV--LGNCKTNKMAPOANLVFQSIMDSXGGLGSPNLQTLFSGQAXSAGAR 330
DB 64 GHTHVAQSV--LGNCKTNKMAPOANLVFQSIMDSXGGLGSPNLQTLFSGQAXSAGAR 330
QY 331 HTNSGAAVNGAYTTDSRNVDDYRKNDMTLLFAAGNEXPN 371
DB 101 -SGQYSWIINGIEMALANNMDVINMSLGPSGSAALKAADVAVAGVVAAGNBSGTS 159
QY 372 G--GTISAPGTAKNAITVGATENLRPSFGSYADNINHAQPSRGPPTKDGRIKPDVAPG 429
DB 160 GSSSTVGYPKYPVIAVGA-----VDSNQRFASFSSVGP-----ELDVMAFG 202
QY 430 TXILSARSSLAPDSSFWANHDSKXYVMGSTMATPIYAGNVA 471
DB 203 VSIQSTLPG-----NKYAGYNGTSMASPHVAGAAA 232
```

```
RESULT 10
US-11-020-602-3
; Sequence 3, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batzell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-020-602-3
```

```
Query Match          7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;
```

```
QY 213 YKADVAQSSYGLYGGQCIYAAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 272
DB 11 IKAP-ALHSQYTGSNVVAVIDSGID-----SSHPLKVAAGSAMVPSRTNPFQDNNS 63
QY 273 GHTHVAQSV--LGNCKTNKMAPOANLVFQSIMDSXGGLGSPNLQTLFSGQAXSAGAR 330
DB 64 GHTHVAQSV--LGNCKTNKMAPOANLVFQSIMDSXGGLGSPNLQTLFSGQAXSAGAR 330
QY 331 HTNSGAAVNGAYTTDSRNVDDYRKNDMTLLFAAGNEXPN 371
DB 101 -SGQYSWIINGIEMALANNMDVINMSLGPSGSAALKAADVAVAGVVAAGNBSGTS 159
QY 372 G--GTISAPGTAKNAITVGATENLRPSFGSYADNINHAQPSRGPPTKDGRIKPDVAPG 429
DB 160 GSSSTVGYPKYPVIAVGA-----VDSNQRFASFSSVGP-----ELDVMAFG 202
```



```

QY      330 IHTNSGAAVNGAYTTDSRVNDDYVRKNDMTILFAAGNEXPNCG--GTISAPGTAKNAITV 387
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      120 VINMSLGCA---SGSTAKMAQAVDNAYARGVYVVAALAGNCGSSGNTTIGTPAKTDSVIAV 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      388 GATENLRPSFGSYADNINHVAQPSRSRPTKGRIRKPDVAPGTXTILSARSSLAPDSSFMA 447
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      177 GA-----VDSNSNRASFSVSG-----AELBWAAPGAGVYSTYPT----- 210
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      448 NHDSKAYMGTSMTPTIYAGNVA-----QLRHFVYKNR 481
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      211 ---NTYATWDTGTSMAFPHVAGAAALLILSKHPNLSASQVNR 248
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

```

US-11-156-062-10
; Sequence 10, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stenif, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-10

```

Query Match 7.2%; Score 221.5; DB 7; Length 274;
 Best Local Similarity 29.9%; Pred. No. 4.2e-10;
 Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

```

QY      212 IYKADVAQSSYGLYQGGQIVAAVADTGLDGRNDSMHEAFRGKITALVALGRTNNANDTN 271
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      10 LIKADKVQAQ-GFKGANVVAVLDTGIGQSHPDLVNVG-----ASFVAGEAYNA-DGN 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      272 GHGTVAGSV--LNGKXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFEOAKSAGAR 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      62 GHGTVAGTVAAALDNTTGVGLVAPSVSYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      330 IHTNSGAAVNGAYTTDSRVNDDYVRKNDMTILFAAGNEXPNCG--GTISAPGTAKNAITV 387
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      120 VINMSLGCA---SGSTAKMAQAVDNAYARGVYVVAALAGNCGSSGNTTIGTPAKTDSVIAV 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      388 GATENLRPSFGSYADNINHVAQPSRSRPTKGRIRKPDVAPGTXTILSARSSLAPDSSFMA 447
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      177 GA-----VDSNSNRASFSVSG-----AELBWAAPGAGVYSTYPT----- 210
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      448 NHDSKAYMGTSMTPTIYAGNVA-----QLRHFVYKNR 481
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      211 ---NTYATWDTGTSMAFPHVAGAAALLILSKHPNLSASQVNR 248
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15
 US-11-156-062-6
 ; Sequence 6, Application US/11156062
 ; Publication No. US20050281773A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wieland, Susanne

```

; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stenif, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-6

```

Query Match 7.2%; Score 220.5; DB 7; Length 274;
 Best Local Similarity 29.9%; Pred. No. 5e-10;
 Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

```

QY      212 IYKADVAQSSYGLYQGGQIVAAVADTGLDGRNDSMHEAFRGKITALVALGRTNNANDTN 271
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      10 LIKADKVQAQ-GFKGANVVAVLDTGIGQSHPDLVNVG-----ASFVAGEAYNA-DGN 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      272 GHGTVAGSV--LNGKXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFEOAKSAGAR 329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      62 GHGTVAGTVAAALDNTTGVGLVAPSVSYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      330 IHTNSGAAVNGAYTTDSRVNDDYVRKNDMTILFAAGNEXPNCG--GTISAPGTAKNAITV 387
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      120 VINMSLGCA---SGSTAKMAQAVDNAYARGVYVVAALAGNCGSSGNTTIGTPAKTDSVIAV 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      388 GATENLRPSFGSYADNINHVAQPSRSRPTKGRIRKPDVAPGTXTILSARSSLAPDSSFMA 447
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      177 GA-----VDSNSNRASFSVSG-----AELBWAAPGAGVYSTYPT----- 210
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      448 NHDSKAYMGTSMTPTIYAGNVA-----QLRHFVYKNR 481
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      211 ---NTYATWDTGTSMAFPHVAGAAALLILSKHPNLSASQVNR 248
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: April 7, 2006, 18:04:19
 Job time : 17.4871 secs

This Page Blank (uspto)

QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGAAXSAGARLHTNSMGAUVNGAYTTDSRNVDVYRKXDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSGAAXSAGARLHTNSMGAUVNGAYTTDSRNVDVYRKXDM 360
QY 361 TILFPAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPPTXG 420
DB 361 TILFPAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPPTXG 420
QY 421 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 421 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 481 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 481 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPXXKXND 600
DB 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPXXKXND 600
QY 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQFSLAIYN 640
DB 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQFSLAIYN 640

RESULT 2

US-10-456-479-4

Sequence 4, Application US/10456479

Publication No. US2004007221A1

GENERAL INFORMATION:

APPLICANT: SATO, TSUYOSHI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: TAKIMURA, YASUHIKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: NOMURA, MASAFUMI
APPLICANT: KOBAYASHI, TOHRU
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 238700USO
CURRENT APPLICATION NUMBER: US/10/456,479
CURRENT FILING DATE: 2003-06-09
PRIOR APPLICATION NUMBER: JP 2002-186387
PRIOR FILING DATE: 2002-06-26
PRIOR APPLICATION NUMBER: JP 2002-304232
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp. KSM-KP43
US-10-456-479-4

Query Match 99.0%; Score 3029; DB 4; Length 640;

Best Local Similarity 93.3%; Pred. No. 8.7e-261;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 MRXKKKVFSLVSAAILSTVALKNPSAGAXXPDLPFGKIQTITDDXXGFSKXQGTGAAA 60
DB 1 MRXKKKVFSLVSAAILSTVALKNPSAGAXXPDLPFGKIQTITDDXXGFSKXQGTGAAA 60
QY 61 FLVSEENVKLLKGLKKLETPANNKLIHQFNGPILBETKQLEXTGAKILDIYIPVAY 120
DB 61 FLVSEENVKLLKGLKKLETPANNKLIHQFNGPILBETKQLEXTGAKILDIYIPVAY 120
QY 121 IVEYEGDVAXKXKXKLEHVESVEPIYPIYXIPDLFTKGAAXIKVXKXALDTQXKKEVDLR 180
DB 121 IVEYEGDVAXKXKXKLEHVESVEPIYPIYXIPDLFTKGAAXIKVXKXALDTQXKKEVDLR 180

QY 181 GIEIXIAOXXXNDVYXYITAKPEYKMDVARGIVKADVAOSSYGLYGGQGLVAAVADTGLD 240
DB 181 GIEIXIAOXXXNDVYXYITAKPEYKMDVARGIVKADVAOSSYGLYGGQGLVAAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGXTNKGMAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGAAXSAGARLHTNSMGAUVNGAYTTDSRNVDVYRKXDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSGAAXSAGARLHTNSMGAUVNGAYTTDSRNVDVYRKXDM 360
QY 361 TILFPAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPPTXG 420
DB 361 TILFPAAGNEXPNVGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQSSRGPPTXG 420
QY 421 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 421 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 481 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 481 RIKPDVMAPGTIIISARSLAPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPXXKXND 600
DB 541 YXFTATAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPXXKXND 600
QY 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQFSLAIYN 640
DB 601 GRNNVENVFINKPOSQTYTIEVQAYNVPGQFSLAIYN 640

RESULT 3

US-10-784-870-6

Sequence 6, Application US/10784870

Publication No. US20040142837A1

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKETI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-10-784-870-6

Query Match 99.0%; Score 3029; DB 4; Length 640;

Best Local Similarity 93.3%; Pred. No. 8.7e-261;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 MRXKKKVFSLVSAAILSTVALKNPSAGAXXPDLPFGKIQTITDDXXGFSKXQGTGAAA 60
DB 1 MRXKKKVFSLVSAAILSTVALKNPSAGAXXPDLPFGKIQTITDDXXGFSKXQGTGAAA 60
QY 61 FLVSEENVKLLKGLKKLETPANNKLIHQFNGPILBETKQLEXTGAKILDIYIPVAY 120

Db 61 FLVESENVKLPKGLQKLETPANNKLIHQFNGPILBETKQOLEKTKAIIIDYIPDYAY 120
Qy 121 IVEYEGDVASXXXXIHEVESVEPYLPYXIDPOLFTKGASXLVKAALDTKQNKKEVQLR 180
Db 121 IVEYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASBLVKAVALDTKQNKKEVQLR 180
Qy 181 GIEIXIQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQGGIYVAVDTGLD 240
Db 181 GIEIXIQFALISNDVXYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQGGIYVAVDTGLD 240
Qy 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNGXTNKGMAPQANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNGXTNKGMAPQANLVF 300
Qy 301 QSIIMDSXGGLGGLPSNLQTLFQOASAGARIHTNSGAAVNGAYTTDSRNVDDYVAKNDM 360
Db 301 QSIIMDSXGGLGGLPSNLQTLFQOASAGARIHTNSGAAVNGAYTTDSRNVDDYVAKNDM 360
Qy 361 TILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Db 361 TILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Qy 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Db 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Qy 481 NRGITPKPSILKALJAGAADKGLGYPNGQMGWVTLDKSLNVAVYNESSXLSSTQKAT 540
Db 481 NRGITPKPSILKALJAGAADKGLGYPNGQMGWVTLDKSLNVAVYNESSXLSSTQKAT 540
Qy 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPXXPKXND 600
Db 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPXXPKXND 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIATV 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIATV 640

RESULT 4
US-10-820-712A-3
Sequence 3, Application US/10820712A
Publication No. US20050026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820, 712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 640
TYPE: PRP
ORGANISM: Bacillus sp. KSM-KP43
US-10-820-712A-3

Query Match 99.0%; Score 3029; DB 5; Length 640;
Best Local Similarity 93.3%; Pred. No. 8.7e-261;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MKKKKVFSLVSAALISTVALXNPSAGXARXFDLPFGKIQTITDDXXGFSKQXGTGAA 60
Db 1 MKKKKVFSLVSAALISTVALXNPSAGXARXFDLPFGKIQTITDDXXGFSKQXGTGAA 60

Qy 61 FLVESENVKXXGKXKGLFTVPANNKLIHQFNGPILBETKQOLEKTKAIIIDYIPDYAY 120
Db 61 FLVESENVKLPKGLQKLETPANNKLIHQFNGPILBETKQOLEKTKAIIIDYIPDYAY 120
Qy 121 IVEYEGDVASXXXXIHEVESVEPYLPYXIDPOLFTKGASXLVKAALDTKQNKKEVQLR 180
Db 121 IVEYEGDVASATSTIEHVESVEPYLPYRIDPOLFTKGASBLVKAVALDTKQNKKEVQLR 180
Qy 181 GIEIXIQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQGGIYVAVDTGLD 240
Db 181 GIEIXIQFALISNDVXYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQGGIYVAVDTGLD 240
Qy 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNGXTNKGMAPQANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNGXTNKGMAPQANLVF 300
Qy 301 QSIIMDSXGGLGGLPSNLQTLFQOASAGARIHTNSGAAVNGAYTTDSRNVDDYVAKNDM 360
Db 301 QSIIMDSXGGLGGLPSNLQTLFQOASAGARIHTNSGAAVNGAYTTDSRNVDDYVAKNDM 360
Qy 361 TILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Db 361 TILFAAGNEXPNGGTSAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGFTKDG 420
Qy 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Db 421 RIKPDVMAFGTILSARSSLPADSSFWANHDSKYAMGTSMAATPIVAGNVAQLRHFYK 480
Qy 481 NRGITPKPSILKALJAGAADKGLGYPNGQMGWVTLDKSLNVAVYNESSXLSSTQKAT 540
Db 481 NRGITPKPSILKALJAGAADKGLGYPNGQMGWVTLDKSLNVAVYNESSXLSSTQKAT 540
Qy 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPXXPKXND 600
Db 541 YKFTATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVYVNDPXXPKXND 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIATV 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVVPVGPQXFSIATV 640

RESULT 5
US-10-820-714A-3
Sequence 3, Application US/10820714A
Publication No. US20050214922A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuoshi
APPLICANT: Kobayashi, Tohru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasushi
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697USO
CURRENT APPLICATION NUMBER: US/10/820, 714A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 640
TYPE: PRP
ORGANISM: Bacillus sp. KSM-KP43
US-10-820-714A-3

Query Match 99.0%; Score 3029; DB 5; Length 640;
Best Local Similarity 93.3%; Pred. No. 8.7e-261;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MKKKKVFSLVSAALISTVALXNPSAGXARXFDLPFGKIQTITDDXXGFSKQXGTGAA 60
Db 1 MKKKKVFSLVSAALISTVALXNPSAGXARXFDLPFGKIQTITDDXXGFSKQXGTGAA 60

SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRF
; ORGANISM: Bacillus sp.
US-10-784-870-8

Query Match 99.0%; Score 3028; DB 4; Length 640;
Best Local Similarity 93.3%; Pred. No. 1,1e-260;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
QY 1 MCKKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTTTDXGFSKXQGTGAAL 60
DB 1 MCKKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTTTAKGFSKXQGTGAAL 60
QY 61 FVSESNVTLKGLKQLETPVANNKLIHQFNGPILBETKQLEXTGAKLIDYIPDYAV 120
DB 61 FVSESNVTLKGLKQLETPVANNKLIHQFNGPILBETKQLEXTGAKLIDYIPDYAV 120
QY 121 IVEEGDVXSKXXKIEHVESVEPYLPXYXIDPOLFTKASXLVKAXALDTKONKEVOLR 180
DB 121 IVEEGDVXSKXXKIEHVESVEPYLPXYXIDPOLFTKASXLVKAXALDTKONKEVOLR 180
QY 181 GIEKIAQXXXSNDVXYITAKPEYKWNDAKGIKADVAQSSYGLYGQGIYAVADTGLD 240
DB 181 GIEQIAQPAISNDVYITAKPEYKWNDAKGIKADVAQSSYGLYGQGIYAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTTHVAGSVLGNXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGTTHVAGSVLGNXTNKGMAPOANLVF 300
QY 301 GSIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDM 360
DB 301 GSIMDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDM 360
QY 361 TILFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSRGGTQKG 420
DB 361 TILFAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSRGGTQKG 420
QY 421 RIKPVMAFGTYILSARSSLAPDSSFWANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 480
DB 421 RIKPVMAFGTYILSARSSLAPDSSFWANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLKAALIAAGADXLGYPNGNQGWRVTLDKSLNVAVYNESSXLSTSQKAT 540
DB 481 NRGITPKPSLKAALIAAGADXLGYPNGNQGWRVTLDKSLNVAVYNESSXLSTSQKAT 540
QY 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXVYVNDPXXPKXXND 600
DB 541 YSFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTQYVYVNDPXXPKXXND 600
QY 601 GRNNVENVFINKPQSGTITIEVOAYNVPGQPSLAIYN 640
DB 601 GRNNVENVFINKPQSGTITIEVOAYNVPGQPSLAIYN 640
```

RESULT 8
US-09-920-954-4
; Sequence 4, Application US/0920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-06-03

PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/J998/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRF
; ORGANISM: Bacillus sp.
US-09-920-954-4

Query Match 98.7%; Score 3020; DB 3; Length 639;
Best Local Similarity 93.4%; Pred. No. 5,5e-260;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
QY 4 KKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTTTDXGFSKXQGTGAAL 63
DB 3 KKKVFLSVLSAAAILSTVALXNPSAGXARXFDLDFKGIQTTTDXGFSKXQGTGAAL 62
QY 64 ESENVTLKGLKQLETPVANNKLIHQFNGPILBETKQLEXTGAKLIDYIPDYAV 123
DB 64 ESENVTLKGLKQLETPVANNKLIHQFNGPILBETKQLEXTGAKLIDYIPDYAV 122
QY 124 YEGDVXSKXXKIEHVESVEPYLPXYXIDPOLFTKASXLVKAXALDTKONKEVOLR 183
DB 124 YEGDVXSKXXKIEHVESVEPYLPXYXIDPOLFTKASXLVKAXALDTKONKEVOLR 182
QY 184 XIAQXXXSNDVXYITAKPEYKWNDAKGIKADVAQSSYGLYGQGIYAVADTGLDTR 243
DB 184 XIAQXXXSNDVXYITAKPEYKWNDAKGIKADVAQSSYGLYGQGIYAVADTGLDTR 242
QY 244 NDSMHEAFRGKITLALYALGRTNNANDTNGHGTTHVAGSVLGNXTNKGMAPOANLVFOSI 303
DB 244 NDSMHEAFRGKITLALYALGRTNNANDTNGHGTTHVAGSVLGNXTNKGMAPOANLVFOSI 302
QY 304 MDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDM 363
DB 304 MDSXGGLGGLPSNLQTLFSQAXSAGARIHTNSWGAAVNGAYTTDSRNVDDYRKNDM 362
QY 364 PAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSRGGTQKG 423
DB 364 PAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSRGGTQKG 422
QY 424 PDVMAFGTYILSARSSLAPDSSFWANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 483
DB 424 PDVMAFGTYILSARSSLAPDSSFWANHDSKTYAMGTSMAATPIVAGNVAQLREHFVK 482
QY 484 ITPKPSLKAALIAAGADXLGYPNGNQGWRVTLDKSLNVAVYNESSXLSTSQKATYX 543
DB 484 ITPKPSLKAALIAAGADXLGYPNGNQGWRVTLDKSLNVAVYNESSXLSTSQKATYX 542
QY 544 TATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXVYVNDPXXPKXXND 603
DB 544 TATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTXVYVNDPXXPKXXND 602
QY 604 NVENVFINKPQSGTITIEVOAYNVPGQPSLAIYN 640
DB 604 NVENVFINKPQSGTITIEVOAYNVPGQPSLAIYN 639
```

RESULT 9
US-10-784-870-4
; Sequence 4, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN


```
; LOCATION: (133)..(133)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (134)..(134)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (147)..(147)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (149)..(149)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (161)..(161)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (184)..(184)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (189)..(189)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (190)..(190)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (195)..(195)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (287)..(287)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (325)..(325)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (370)..(370)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (502)..(502)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (532)..(532)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (542)..(542)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (585)..(585)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (592)..(592)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (593)..(593)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (595)..(595)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (596)..(596)
```

```
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (597)..(597)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (612)..(612)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (633)..(633)
; OTHER INFORMATION: Xaa is any amino acid
; US-09-920-954-2
```

Query Match 98.6%; Score 3016; DB 3; Length 640;
Best Local Similarity 100.0%; Pred. No. 1,3e-259; Indels 0; Gaps 0;
Matches 640; Conservative 0; Mismatches 0;

```
QY 1 MEXKKKVFSLVLSAAAILSTVALXNPASAGAXXFDLDFGIQTDDXKGFSSKXOTGAHA 60
|
|
|
Db 1 MEXKKKVFSLVLSAAAILSTVALXNPASAGAXXFDLDFGIQTDDXKGFSSKXOTGAHA 60
|
|
|
QY 61 FLVSEENVKLXKGLKKLETPANNKGLHXQFNGLIESTKXKLEXTGAKIIDIYIPVAY 120
|
|
|
Db 61 FLVSEENVKLXKGLKKLETPANNKGLHXQFNGLIESTKXKLEXTGAKIIDIYIPVAY 120
|
|
|
QY 121 IVEYEGDVAXSXXXIIEHVESVEPYLPXYXIDPOLPTKGASXLVKXALDPTKXNKEVQLR 180
|
|
|
Db 121 IVEYEGDVAXSXXXIIEHVESVEPYLPXYXIDPOLPTKGASXLVKXALDPTKXNKEVQLR 180
|
|
|
QY 181 GIEIXIAQXXXSNDVYYITAKPEYKVMNDVARGIVRADVQSSGYGGQGIYVAVDTGLD 240
|
|
|
Db 181 GIEIXIAQXXXSNDVYYITAKPEYKVMNDVARGIVRADVQSSGYGGQGIYVAVDTGLD 240
|
|
|
QY 241 TGRNDSNMHEAFRGKITLVALGRTNNANDTNGHGHVAGSVLGGXCTNKGAPQANLVF 300
|
|
|
Db 241 TGRNDSNMHEAFRGKITLVALGRTNNANDTNGHGHVAGSVLGGXCTNKGAPQANLVF 300
|
|
|
QY 301 QSINMSXGGLGCLPBNLQTLFQOAXSAGARHTTNSWGAAVNCAYTTDSANNVDYRKNDM 360
|
|
|
Db 301 QSINMSXGGLGCLPBNLQTLFQOAXSAGARHTTNSWGAAVNCAYTTDSANNVDYRKNDM 360
|
|
|
QY 361 TILPAAGNEXPMGCTISAPGTAKNAITVGATENLSPFSGYADNINHVAQFSSRGPTKDG 420
|
|
|
Db 361 TILPAAGNEXPMGCTISAPGTAKNAITVGATENLSPFSGYADNINHVAQFSSRGPTKDG 420
|
|
|
QY 421 RIKPDVMAPGTYILSARSLAPDSSFMANHDSKAYMGSTNATPIVAGNVAQLEHFFVK 480
|
|
|
Db 421 RIKPDVMAPGTYILSARSLAPDSSFMANHDSKAYMGSTNATPIVAGNVAQLEHFFVK 480
|
|
|
QY 481 NRGITPKPSLKAALIAAADXGLGYPNGNCGWGYTLDKSLNVAIVNNESSXLSTSOKAT 540
|
|
|
Db 481 NRGITPKPSLKAALIAAADXGLGYPNGNCGWGYTLDKSLNVAIVNNESSXLSTSOKAT 540
|
|
|
QY 541 YXFATAGKPLKISLWSDPASTTASVTLVNDLVLVTAPNGTYVNGDFPXXPKXNDM 600
|
|
|
Db 541 YXFATAGKPLKISLWSDPASTTASVTLVNDLVLVTAPNGTYVNGDFPXXPKXNDM 600
|
|
|
QY 601 GRNNVENYFINXPOSGTYTIEVOAYNVVPGQFSLAIYN 640
|
|
|
Db 601 GRNNVENYFINXPOSGTYTIEVOAYNVVPGQFSLAIYN 640
|
|
|
RESULT 11
; US-10-784-870-2
; Sequence 2, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
```

APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid

FEATURE:
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:


```
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-2
```

```
Query Match 98.6%; Score 3016; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 1,3e-259;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRXKKKVLSTLSAAIISTVALXNPSAGXAXFPDLDPFGIOTTTDXKGFSGKOTGAAA 60
DB 1 MRXKKKVLSTLSAAIISTVALXNPSAGXAXFPDLDPFGIOTTTDXKGFSGKOTGAAA 60
QY 61 FLVESBENVKLGKGLKKTETVPANNKLMHXOPNGPILBETKQXLEXTGAKLIDYIPDAY 120
DB 61 FLVESBENVKLGKGLKKTETVPANNKLMHXOPNGPILBETKQXLEXTGAKLIDYIPDAY 120
QY 121 IVEYEGDVYKXXXXIEHVESVEPYLPXYXIDPQLFTKSGSLVKXALDTRKXNREVQLR 180
DB 121 IVEYEGDVYKXXXXIEHVESVEPYLPXYXIDPQLFTKSGSLVKXALDTRKXNREVQLR 180
QY 181 GIEKTAQXXXXNDVYITRAKPEYKVMNDVARGIVADVAQSSYGLYGQQLVAVADTGID 240
DB 181 GIEKTAQXXXXNDVYITRAKPEYKVMNDVARGIVADVAQSSYGLYGQQLVAVADTGID 240
QY 241 TGRNDSMHEAPRGTITLALGRTNNANDTNGHGVAGSYLGNKXTNKGAPQANLVF 300
DB 241 TGRNDSMHEAPRGTITLALGRTNNANDTNGHGVAGSYLGNKXTNKGAPQANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHITNSWGAAVNGAYTTDSRNVDDYVRKIDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSGQASAGARIHITNSWGAAVNGAYTTDSRNVDDYVRKIDM 360
```

```
QY 361 TLTPAAGNEXPNGGTITAPGTAKNAITVGATENTLAPSPGCVADNINHVAQFSSRGPTKDG 420
DB 361 TLTPAAGNEXPNGGTITAPGTAKNAITVGATENTLAPSPGCVADNINHVAQFSSRGPTKDG 420
QY 421 RIKPDVMAPGTYILSARSLAPDSSFWANHDSKAYMGTSNATPIVAGNVQLREHPYK 480
DB 421 RIKPDVMAPGTYILSARSLAPDSSFWANHDSKAYMGTSNATPIVAGNVQLREHPYK 480
QY 481 NRGITPKSLLKALILAGAADXGLGYPNGOQGWAVTLDKSLINVAVYNESSXLSTSOKAT 540
DB 481 NRGITPKSLLKALILAGAADXGLGYPNGOQGWAVTLDKSLINVAVYNESSXLSTSOKAT 540
QY 541 YXFTATKRPKLKLSIWSDBAPASTASTLVNDLVLVTTAPNGTYTVGDPFXKPPXXNMD 600
DB 541 YXFTATKRPKLKLSIWSDBAPASTASTLVNDLVLVTTAPNGTYTVGDPFXKPPXXNMD 600
QY 601 GRNNVENVFINKPQSGTYTIEVQAVNVPVGPQXPSLAIYN 640
DB 601 GRNNVENVFINKPQSGTYTIEVQAVNVPVGPQXPSLAIYN 640

RESULT 12
US-09-920-954-1
Sequence 1, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPR: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
```

LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
us-09-920-954-1

Query Match 98.3%; Score 3007; DB 3; Length 639;
Best Local Similarity 100.0%; Pred. No. 7.9e-259;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 KKKVFLVLSAAALSTVALXNPSAGXARXFDLDFKGIOTTTDXKXGSKOXQTAAAFV 63
3 KKKVFLVLSAAALSTVALXNPSAGXARXFDLDFKGIOTTTDXKXGSKOXQTAAAFV 62
64 ESENVKLXKGLXKKLETPVANNKLIHQFNGPILEFTKXKLETKGALIDYIPYAYIVE 123
63 ESENVKLXKGLXKKLETPVANNKLIHQFNGPILEFTKXKLETKGALIDYIPYAYIVE 122
124 YEGDVXSKXXKIEHVESVEPYLPYXIDPOLFTKGSXLVKAALDTKXNKEVQLGIE 183
123 YEGDVXSKXXKIEHVESVEPYLPYXIDPOLFTKGSXLVKAALDTKXNKEVQLGIE 182
184 XIAQXXSNVYXYTTAKPEKVMNDVARGIVKADVAQSGVLGGQGI VAVADTGLDGR 243
183 XIAQXXSNVYXYTTAKPEKVMNDVARGIVKADVAQSGVLGGQGI VAVADTGLDGR 242
244 NDSMHBAPFGKITLALYALGRTNANDTNGHGVAGSVLGNXTNGMAPOANLVFQSI 303
243 NDSMHBAPFGKITLALYALGRTNANDTNGHGVAGSVLGNXTNGMAPOANLVFQSI 302
304 MDSYGLGGLPSNIQTLFSDXSKAGARIHTNSGAAVNGAYTTDSRVDDYVRKNDMTIL 363
303 MDSYGLGGLPSNIQTLFSDXSKAGARIHTNSGAAVNGAYTTDSRVDDYVRKNDMTIL 362
364 PAAGNEXPNGGTISAPETAANAITYGATEMLRPSFGSYANINHVQOFSSRGPTKGRIR 423
363 PAAGNEXPNGGTISAPETAANAITYGATEMLRPSFGSYANINHVQOFSSRGPTKGRIR 422
424 PDVWAPGTXTILSARSSILAPDSSFMANHDSKYAVNGTSMATPIVAGNVALRHFVNRG 483
423 PDVWAPGTXTILSARSSILAPDSSFMANHDSKYAVNGTSMATPIVAGNVALRHFVNRG 482
484 ITPKPSILKAAALTAGAADXGLGYPNGNGGVRVTLDKSLNVAAYNBSSXLSTSQKATYX 543
483 ITPKPSILKAAALTAGAADXGLGYPNGNGGVRVTLDKSLNVAAYNBSSXLSTSQKATYX 542

Qy	544	TATAGKPLKSLVWSDAPASTASTATLVNDLVTATPAGTYVNGDPFKKPPXXNWDGN	603	OTHER INFORMATION: Xaa is any amino acid
Db	543	TATAGKPLKSLVWSDAPASTASTATLVNDLVTATPAGTYVNGDPFKKPPXXNWDGN	602	FEATURE: NAME/KEY: misc_feature LOCATION: (102)..(102) OTHER INFORMATION: Xaa is any amino acid
Qy	604	NVENVFINKPQSGTTTIEVQAYNVFVGPQXPSLALVN	640	FEATURE: NAME/KEY: misc_feature LOCATION: (105)..(105) OTHER INFORMATION: Xaa is any amino acid
Db	603	NVENVFINKPQSGTTTIEVQAYNVFVGPQXPSLALVN	639	FEATURE: NAME/KEY: misc_feature LOCATION: (128)..(128) OTHER INFORMATION: Xaa is any amino acid
RESULT 13				
US-10-784-870-1				
Sequence 1, Application US/10784870				
Publication No. US20040142837A1				
GENERAL INFORMATION:				
APPLICANT: TAKAIWA, MIKIO				
APPLICANT: OKUDU, MITSUYOSHI				
APPLICANT: SAKKI, KATSUYO				
APPLICANT: KUBOTA, HIROMI				
APPLICANT: HITOMI, JUN				
APPLICANT: KAGEYAMA, YASUSHI				
APPLICANT: SHIKATA, SHITSUM				
APPLICANT: NOMURA, MASAFUMI				
TITLE OF INVENTION: ALKALINE PROTEASE				
FILE REFERENCE: 0327-0832-0PCT				
CURRENT APPLICATION NUMBER: US/10/784,870				
CURRENT FILING DATE: 2004-02-24				
PRIOR APPLICATION NUMBER: US/09/509,814A				
PRIOR FILING DATE: 2000-04-06				
PRIOR APPLICATION NUMBER: PCT/JP98/04528				
PRIOR FILING DATE: 1998-10-07				
PRIOR APPLICATION NUMBER: JP 9-274570				
PRIOR FILING DATE: 1997-06-08				
NUMBER OF SEQ ID NOS: 24				
SOFTWARE: PatentIn version 3.0				
SEQ ID NO 1				
LENGTH: 639				
TYPE: PRT				
ORGANISM: Bacillus sp.				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (123)..(123)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (129)..(129)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (132)..(132)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (146)..(146)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (147)..(147)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (153)..(153)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (170)..(170)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (174)..(174)				
OTHER INFORMATION: Xaa is any amino acid				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (189)..(189)				

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-10-784-870-1
```

```
Query Match 98.3%; Score 3007; DB 4; Length 639;
Best Local Similarity 100.0%; Pred. No. 7.9e-259;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 KKKVFLSVLSAAALISTVALNPSAGXAXFDLPFGKIQTITDDXGFSKXQOTGAAPLV 63
DB 3 KKKVFLSVLSAAALISTVALNPSAGXAXFDLPFGKIQTITDDXGFSKXQOTGAAPLV 62
QY 64 ESEVVKLKKGLKKGLKLETPANNKLIHXQFNGPILEETXQILETGAKTLDYIPYAYIVE 123
DB 63 ESEVVKLKKGLKKGLKLETPANNKLIHXQFNGPILEETXQILETGAKTLDYIPYAYIVE 122
```

```
QY 124 YEGDVXSKXXIIHEVESVEPLPYXIIDPOLFTKGASXLVKAALDTKONKKEQVLAIGIE 183
DB 123 YEGDVXSKXXIIHEVESVEPLPYXIIDPOLFTKGASXLVKAALDTKONKKEQVLAIGIE 182
QY 184 XIAQXXXNDVXYITTAPEYKVMNDVARGIYKADVAOSYGLYGQGOI VAVADTGLDTGR 243
DB 183 XIAQXXXNDVXYITTAPEYKVMNDVARGIYKADVAOSYGLYGQGOI VAVADTGLDTGR 242
QY 244 NDSMHEAFRGKITALYALGRTNANDTNGHGTIVAGSVLGNXTNKMAPOANLVFQSI 303
DB 243 NDSMHEAFRGKITALYALGRTNANDTNGHGTIVAGSVLGNXTNKMAPOANLVFQSI 302
QY 304 MDSYGGIGGLPSNLTQLPSQAXSAGARIHNSWGAANVAGATYDTSRVVDVVRKNDWTIL 363
DB 303 MDSYGGIGGLPSNLTQLPSQAXSAGARIHNSWGAANVAGATYDTSRVVDVVRKNDWTIL 362
QY 364 PAAGNEXPNGGTISAPGTAKNAITVGATENLRPFSGSYADININVAQFSSRGPTKGRIR 423
DB 363 PAAGNEXPNGGTISAPGTAKNAITVGATENLRPFSGSYADININVAQFSSRGPTKGRIR 422
QY 424 PDVAPGTXTIISARSSIAPDSSFWANHDSKYAVWGTSMATPIVAGNVAQLREHFVXNRG 483
DB 423 PDVAPGTXTIISARSSIAPDSSFWANHDSKYAVWGTSMATPIVAGNVAQLREHFVXNRG 482
QY 484 ITPKPSLLKALITGADXGIGYPNGOGNGRVTLDLSLVAVYNESSXISTGKATYXP 543
DB 483 ITPKPSLLKALITGADXGIGYPNGOGNGRVTLDLSLVAVYNESSXISTGKATYXP 542
QY 544 TATAGKPLKISLVWSDAPASTTASVTLVNDLVTAPNGTXVYVGNDFXXPXXKNMGRN 603
DB 543 TATAGKPLKISLVWSDAPASTTASVTLVNDLVTAPNGTXVYVGNDFXXPXXKNMGRN 602
QY 604 NVENVFINXPOSGTYTTEVOAIVNVPVGPQXSLAIYN 640
DB 603 NVENVFINXPOSGTYTTEVOAIVNVPVGPQXSLAIYN 639
```

```
RESULT 14
US-09-985-689A-1
; Sequence 1, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGWAKI, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-1
```

```
Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. No. 4e-183;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
```

```
QY 207 NDVARGIVKADVAVQSSYGLYGQGOI VAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
```

Db 1 NDVARGIVADVAQSSYGLYGQGIIVAVADTGLDGRNDSMHBAFRGKITLALYALGRTN 60
Qy 267 NANDNMGHTHVAAGTGLGNGXTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFEOAXS 326
Db 61 NANDNMGHTHVAAGTGLGNGXTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFEOAXS 120
Qy 327 AGARITHNSMGAAVNGATTTDSRNVDYRKNDMTILFAAGNEXPGGTISAPGTAKNAI 386
Db 121 AGARITHNSMGAAVNGATTTDSRNVDYRKNDMTILFAAGNEXPGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAPGTYILSARSSLABDSF 446
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAPGTYILSARSSLABDSF 240
Qy 447 WANHSKAYMGSTMAATPIVAGNVAQLEHFPVKRGITPKPSLLKALIIAGAADXGLGY 506
Db 241 WANHSKAYMGSTMAATPIVAGNVAQLEHFPVKRGITPKPSLLKALIIAGAADXGLGY 300
Qy 507 PNGOQMGKRVTLDKSLNVAAYNNESSXLSTOKATYXFTTAGKPLKISLWSDAPASTTA 566
Db 301 PNGOQMGKRVTLDKSLNVAAYNNESSXLSTOKATYXFTTAGKPLKISLWSDAPASTTA 360
Qy 567 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 626
Db 361 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 420
Qy 627 VPVGPOXFSLAIYN 640
Db 421 VPVGPOXFSLAIYN 434

RESULT 15

US-09-985-689A-2
; Sequence 2, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 3; Length 434;
Best Local Similarity 96.3%; Pred. No. 4e-183;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 207 NDVARGIVADVAQSSYGLYGQGIIVAVADTGLDGRNDSMHBAFRGKITLALYALGRTN 266
Db 1 NDVARGIVADVAQSSYGLYGQGIIVAVADTGLDGRNDSMHBAFRGKITLALYALGRTN 60
Qy 267 NANDNMGHTHVAAGTGLGNGXTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFEOAXS 326
Db 61 NANDNMGHTHVAAGTGLGNGXTTKGMAPOANLVFQSIMDSXGGLGGLPSNLQTLFEOAXS 120
Qy 327 AGARITHNSMGAAVNGATTTDSRNVDYRKNDMTILFAAGNEXPGGTISAPGTAKNAI 386

Db 121 AGARITHNSMGAAVNGATTTDSRNVDYRKNDMTILFAAGNEXPGGTISAPGTAKNAI 180
Qy 387 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAPGTYILSARSSLABDSF 446
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKPDVMAPGTYILSARSSLABDSF 240
Qy 447 WANHSKAYMGSTMAATPIVAGNVAQLEHFPVKRGITPKPSLLKALIIAGAADXGLGY 506
Db 241 WANHSKAYMGSTMAATPIVAGNVAQLEHFPVKRGITPKPSLLKALIIAGAADXGLGY 300
Qy 507 PNGOQMGKRVTLDKSLNVAAYNNESSXLSTOKATYXFTTAGKPLKISLWSDAPASTTA 566
Db 301 PNGOQMGKRVTLDKSLNVAAYNNESSXLSTOKATYXFTTAGKPLKISLWSDAPASTTA 360
Qy 567 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 626
Db 361 SVTLVNDLDELVTAPNGTXYVGNDFXKPKXXNMDGRNNEVFINXPOSGTYTIEVOAYN 420
Qy 627 VPVGPOXFSLAIYN 640
Db 421 VPVGPOXFSLAIYN 434

Search completed: April 7, 2006, 18:03:41
Job time : 119.593 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:32:32 ; Search time 138.608 Seconds
(without alignments)
2028.756 Million cell updates/sec

Title: US-10-784-870-2
Perfect score: 3059
Sequence: 1 MRXXXXVFLSLASAAIIST.....EVQAVNPVGPQPSLAIYN 640

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3029	99.0	640	AA117090	AA117090 Bacillus
2	3029	99.0	640	ADM40773	ADM40773 Alkaline
3	3029	99.0	640	ADSI4437	ADSI4437 Bacillus
4	3028	99.0	641	ADT49613	ADT49613 Bacillus
5	3028	99.0	640	AA117091	AA117091 Bacillus
6	3020	98.7	639	AA117089	AA117089 Bacillus
7	3016	98.6	640	AA117088	AA117088 An alkali
8	3007	98.3	639	AA117087	AA117087 An alkali
9	2722.5	89.0	641	AAW89547	AAW89547 Bacillus
10	2414	78.9	636	AAW89548	AAW89548 Bacillus
11	2155	70.4	434	AAW50080	AAW50080 Bacillus
12	2155	70.4	434	AAW50081	AAW50081 Bacillus
13	2155	70.4	434	ADY33778	ADY33778 Bacillus
14	2155	70.4	434	ADZ51758	ADZ51758 Mutant Ba
15	2155	70.4	434	ADZ51757	ADZ51757 Wild-type
16	2155	70.4	434	ADL25802	ADL25802 Bacillus
17	2155	70.4	434	ADM40779	ADM40779 Alkaline
18	2155	70.4	434	ADM40771	ADM40771 Mature al
19	2155	70.4	434	ADM40780	ADM40780 Alkaline
20	2155	70.4	434	ADSI4427	ADSI4427 Bacillus
21	2155	70.4	434	ADT49604	ADT49604 Bacillus
22	2147	70.2	434	ADSI4441	ADSI4441 Bacillus
23	2147	70.2	434	ADSI4438	ADSI4438 Bacillus
24	2147	70.2	434	ADSI4442	ADSI4442 Bacillus

25	2146	70.2	434	8	ADSI4439	ADSI4439 Bacillus
26	2140	70.0	434	8	ADSI4440	ADSI4440 Bacillus
27	2138	69.9	434	8	ADSI4443	ADSI4443 Bacillus
28	2133	69.7	434	8	ADSI4444	ADSI4444 Bacillus
29	2118	69.2	434	8	ADSI4445	ADSI4445 Alkaline
30	2103	68.7	436	8	ADM40787	ADM40787 Bacillus
31	2082	68.1	434	5	AAW50085	AAW50085 Bacillus
32	2082	68.1	434	7	ADZ51762	ADZ51762 Mutant Ba
33	2082	68.1	434	8	ADM40784	ADM40784 Alkaline
34	2064.5	67.5	433	8	ADSI52082	ADSI52082 Bacillus
35	2061.5	67.4	433	8	ADSI52083	ADSI52083 Bacillus
36	2060.5	67.4	433	5	AAW50086	AAW50086 Bacillus
37	2060.5	67.4	433	7	ADZ51763	ADZ51763 Mutant Ba
38	2060.5	67.4	433	8	ADM40785	ADM40785 Alkaline
39	2060.5	67.4	433	8	ADSI52010	ADSI52010 Bacillus
40	2060.5	67.4	433	8	ADSI52016	ADSI52016 Bacillus
41	2060.5	67.4	433	8	ADSI52011	ADSI52011 Bacillus
42	2060.5	67.4	433	8	ADSI52001	ADSI52001 Bacillus
43	2059.5	67.3	433	8	ADSI52062	ADSI52062 Bacillus
44	2059.5	67.3	433	8	ADSI52054	ADSI52054 Bacillus
45	2058.5	67.3	433	8	ADSI52054	ADSI52054 Bacillus

ALIGNMENTS

RESULT 1	AA117090	standard; protein; 640 AA.
XX	AA117090	
AC	AA117090;	
DT	20-MAR-2003 (revised)	
DT	21-JUL-1999 (first entry)	
XX		
DE	Bacillus alkaline protease.	
XX		
KW	Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;	
KW	washing composition; oxidising agent.	
XX		
OS	Bacillus sp.	
XX		
PN	MO9918218-A1.	
PD	15-APR-1999.	
XX		
PF	07-OCT-1998; 98WO-JP004528.	
XX		
PR	07-OCT-1997; 97JP-00274570.	
XX		
PA	(KAOS) KAO CORP.	
PI	Takeiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;	
PI	Shikata S, Nomura M,	
XX	WPI: 1999-287736/27.	
DR	N-PSDB; AAX37278.	
XX		
PT	Alkali protease from Bacillus used in washing powders.	
XX		
PS	Disclosure: Page 58-63; 71pp; Japanese.	
XX		
CC	The invention relates to alkaline proteases produced by strains of	
CC	Bacillus. The proteases ability to digest casein is not inhibited by	
CC	oleic acid and they have a high stability to oxidising agents. The	
CC	alkaline protease of the invention has the following properties: (a) It	
CC	is active over the pH range 4-13 and has at least 80% of its optimum	
CC	activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is	
CC	stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)	
CC	its ability to digest casein is not inhibited by oleic acid; (e) it has	
CC	molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be	
CC	used as enzymes in washing compositions for use in automatic dishwashers	
CC	and for washing clothes. The stability to oxidising agents allows the	

CC enzyme to be an effective component of washing compositions including
CC bleachers. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)

CC Sequence 640 AA;

Query Match 99.0%; Score 3029; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.6e-237;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
QY 1 MRKKKKVFLSVLSAAAILSTVALXNPSAGAXRFDLPFKGIQTITDXXGFSKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALXNPSAGAXRFDLPFKGIQTITDXXGFSKQGTGAAA 60
QY 61 FLVSEENVKLGKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 120
DB 61 FLVSEENVKLGKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 120
QY 121 IVEYEGDVXSKXXXIEHVESVEPYLPYXIDPQLFTKGASXLVKAALDTKXNKEVQLR 180
DB 121 IVEYEGDVXSKATSTIEHVESVEPYLPYRIDPQLFTKGASBLVKAALDTKXNKEVQLR 180
QY 181 GIEKIAQXXXSNDVXYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDTGLD 240
DB 181 GIEQIAQFAISNDVLYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNKXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNKXTNKGMAPOANLVF 300
QY 301 OSIINDSGGLGGLPSNLQTLFSGQXASAGARIHTNSWGAAVNGAYTTDSRVDDVYRKNDM 360
DB 301 OSIINDSGGLGGLPSNLQTLFSGQXASAGARIHTNSWGAAVNGAYTTDSRVDDVYRKNDM 360
QY 361 TILFAAGNEXPNGTITISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
DB 361 TILFAAGNEXPNGTITISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
QY 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLLKALIIAGAADXGLGYPNGNGMGWVTLDKSLNAVYNSSSLSTSQKAT 540
DB 481 NRGITPKPSLLKALIIAGAADXGLGYPNGNGMGWVTLDKSLNAVYNSSSLSTSQKAT 540
QY 541 YKFTATACKPLKISLWSDAPASTTASVTLVNDLDTLITAPNGTXYVGNDFXPPKXNMD 600
DB 541 YKFTATACKPLKISLWSDAPASTTASVTLVNDLDTLITAPNGTXYVGNDFXPPKXNMD 600
QY 601 GRNNVENVFINXPOSQGTITIEVOAYNVVPGQPSLAIVN 640
DB 601 GRNNVENVFINAPOSQGTITIEVOAYNVVPGQPSLAIVN 640
```

RESULT 2
ADM40773

ID ADM40773 standard; protein; 640 AA.

XX AC ADM40773;

XX DT 01-JUL-2004 (first entry)

XX DE Alkaline protease from *Bacillus* sp. KSM-KP43.

XX KW alkaline protease; laundry detergent; bleaching agent; detergent;

XX OS denture-cleaning agent; enzyme.

XX XX *Bacillus* sp.; KSM-KP43.

XX PN US2004072321-A1.

XX PD 15-APR-2004.

XX 09-JUN-2003; 2003US-00456479.

XX 26-JUN-2002; 2002JP-00186387.

XX 18-OCT-2002; 2002JP-00304232.

XX PA (KAO) KAO CORP.

XX PI Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;

XX DR MPI; 2004-328572/30.

XX N-PSDB; ADM40772.

PT New alkaline protease having a mutant prepro sequence where amino acid
PT residues at positions 52, 75 and 142 are substituted with another amino
PT acid residue, useful as enzyme component of laundry detergents, or
PT bleaching agents.

XX PS Disclosure; SEQ ID NO 4; 29pp; English.

CC The invention relates to an alkaline protease having a prepro sequence.

CC The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
CC acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
CC acid residues at: (a) position 52 is substituted by aspartic acid or
CC arginine; (b) position 75 is substituted by alanine or arginine; and (c)
CC position 142 is substituted by lysine. The alkaline protease is useful as
CC an enzyme which can be incorporated into laundry detergents, bleaching
CC agents, detergents for cleaning hard surfaces or drainpipes, denture-
CC cleaning agents, and detergents for sterilizing medical apparatus. The
CC present sequence represents alkaline protease from *Bacillus* sp. KSM-KP43.

XX Sequence 640 AA;

Query Match 99.0%; Score 3029; DB 8; Length 640;
Best Local Similarity 93.3%; Pred. No. 1.6e-237;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
QY 1 MRKKKKVFLSVLSAAAILSTVALXNPSAGAXRFDLPFKGIQTITDXXGFSKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALXNPSAGAXRFDLPFKGIQTITDXXGFSKQGTGAAA 60
QY 61 FLVSEENVKLGKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 120
DB 61 FLVSEENVKLGKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 120
QY 121 IVEYEGDVXSKXXXIEHVESVEPYLPYXIDPQLFTKGASXLVKAALDTKXNKEVQLR 180
DB 121 IVEYEGDVXSKATSTIEHVESVEPYLPYRIDPQLFTKGASBLVKAALDTKXNKEVQLR 180
QY 181 GIEKIAQXXXSNDVXYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDTGLD 240
DB 181 GIEQIAQFAISNDVLYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQIVAAVDTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNKXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRNNANDTNGHGTIVAGSYLGNKXTNKGMAPOANLVF 300
QY 301 OSIINDSGGLGGLPSNLQTLFSGQXASAGARIHTNSWGAAVNGAYTTDSRVDDVYRKNDM 360
DB 301 OSIINDSGGLGGLPSNLQTLFSGQXASAGARIHTNSWGAAVNGAYTTDSRVDDVYRKNDM 360
QY 361 TILFAAGNEXPNGTITISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
DB 361 TILFAAGNEXPNGTITISAPGTAKNAITVGATENLRPSFGSYADININVAQSSRGPTKDG 420
QY 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKAYMGSTSMATPIVAGNVAQLREHFVK 480
DB 421 RIKPDVWAPGTXYILSARSSSLAPDSSFANNDSKAYMGSTSMATPIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLLKALIIAGAADXGLGYPNGNGMGWVTLDKSLNAVYNSSSLSTSQKAT 540
DB 481 NRGITPKPSLLKALIIAGAADXGLGYPNGNGMGWVTLDKSLNAVYNSSSLSTSQKAT 540
```


Qy 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXVYVNDPXXKXND 600
 Db 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXVYVNDPXXKXND 600
 Qy 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSIAIYN 640
 Db 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSIAIYN 640

RESULT 3

ADSI4437
 ID ADSI4437 standard; protein; 640 AA.

AC ADSI4437;

DT 30-DEC-2004 (first entry)

DE Bacillus alkaline protease KP43.

KW protease; enzyme; alkaline protease; laundry detergent; KP43.

OS Bacillus sp.; KSM-KP43.

FX Key Location/Qualifiers

FT Peptide 1..206

FT Protein 207..640

PN EPI466962-A1.

PD 13-OCT-2004.

PF 08-APR-2004; 2004BP-00008604.

PR 10-APR-2003; 2003JP-00106709.

PA (KAOS) KAO CORP.

PI Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;

DR WPI, 2004-711317/70.

DR N-PSDB; ADSI4428.

PT New engineered alkaline protease, useful particularly in laundry

PS detergents, comprising specified amino acids at particular positions.

PS Disclosure; SEQ ID NO 1; 31pp; English.

CC The invention relates to a novel alkaline protease. The new alkaline

CC protease comprises an amino acid sequence in which one or more amino acid

CC ADSI4427, or at positions corresponding to these positions are: position

CC 15 (histidine), position 16 (threonine or glutamine), position 166

CC (glycine), position 167 (valine), position 187 (serine), position 166

CC (arginine), and position 405 (aspartic acid). The alkaline protease is

CC useful in industry particularly in laundry detergents, but also e.g. in

CC fibre modifying agents, leather processing agents, cosmetic compositions,

CC bath additives, food-modifying agents, and pharmaceuticals. The present

CC sequence represents the wild-type Bacillus sp. KSM-KP43 alkaline

CC protease, KP43. The sequence is shown in the sequence listing as part of

CC SEQ ID NO:1

CC Sequence 640 AA;

Query Match 99.0%; Score 3029; DB 8; Length 640;

Best Local Similarity 93.3%; Pred. No. 1.6e-237;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MKKKKVPISVLSAAIISTVALNPSAGAKXPDLDGFIQTDDXXKXQXGAA 60

Db 1 MKKKKVPISVLSAAIISTVALNPSAGAKXPDLDGFIQTDDXXKXQXGAA 60

Qy 61 FLVSEBNVGLKGLKKTETVPANNKGLHIXQFNGPILEETKQXLEXTGAKLIDYIPVAY 120

Db 61 FLVSEBNVGLKGLKKTETVPANNKGLHIXQFNGPILEETKQXLEXTGAKLIDYIPVAY 120
 Qy 121 IYVEGADVSKXXXIEHVSEVEBYPXYYIDPOLFTKGSXIVKANALDTKXNKEVQLR 180
 Db 121 IYVEGADVSKATSTIEHVSEVEBYPXYYIDPOLFTKGSXIVKANALDTKXNKEVQLR 180
 Qy 181 GIEIXIOXXXSNVYXITAKPEYKMNVDVARGIVKADVQVSGVYGQGIYVAVDPTGLD 240
 Db 181 GIEIXIOXXXSNVYXITAKPEYKMNVDVARGIVKADVQVSGVYGQGIYVAVDPTGLD 240
 Qy 241 TGRNDSMHEAFKGTITLALYALGRTNNADDTNGHGTGVGSVVGNGXTNKGMAPOANLVP 300
 Db 241 TGRNDSMHEAFKGTITLALYALGRTNNADDTNGHGTGVGSVVGNGXTNKGMAPOANLVP 300
 Qy 301 QSIMDSXGIGLPSNLQTLFSGVXSGARIHITNSGAAVNGAYTTDSRVVDVYKNDM 360
 Db 301 QSIMDSGGLGGLPSNLQTLFSGVXSGARIHITNSGAAVNGAYTTDSRVVDVYKNDM 360
 Qy 361 TLIFPAGNEXPNCGTISAPGTAKNALTVGATENTLRPSFSGYADNINHVAFSSRGPTKOG 420
 Db 361 TLIFPAGNEXPNCGTISAPGTAKNALTVGATENTLRPSFSGYADNINHVAFSSRGPTKOG 420
 Qy 421 RIKPDVMAFGTYILSARSSLAPDSFMANHDSKYAVMGTSMATPIVAGVNAQLRHFYK 480
 Db 421 RIKPDVMAFGTYILSARSSLAPDSFMANHDSKYAVMGTSMATPIVAGVNAQLRHFYK 480
 Qy 481 NRGITPKPSLLKALIAAGADYGLGYPNGNGMGRVTLDKSLNVAVYNESSXLSIQKAT 540
 Db 481 NRGITPKPSLLKALIAAGADYGLGYPNGNGMGRVTLDKSLNVAVYNESSXLSIQKAT 540
 Qy 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXVYVNDPXXKXND 600
 Db 541 YKFTATAGKPLKLSLVWSDAPASTTASVTLVNDLDTVITAPNGTXVYVNDPXXKXND 600
 Qy 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSIAIYN 640
 Db 601 GRNNVENVPINXPOSGTYYTIEVOAYNVPGPOXFSIAIYN 640

RESULT 4

ID ADT49613 standard; protein; 641 AA.

AC ADT49613;

DT 30-DEC-2004 (first entry)

DE Bacillus alkaline protease.

KW Alkaline protease; detergent; fiber modification; leather processing;

KW cosmetic; bath additives; food-modification; pharmaceutical; enzyme.

OS Bacillus sp. KSM-KP43.

FX Key Location/Qualifiers

FT Peptide 1..206

FT Protein 207..641

PN EPI466970-A1.

PD 13-OCT-2004.

PF 08-APR-2004; 2004BP-00008605.

PR 10-APR-2003; 2003JP-00106708.

PA (KAOS) KAO CORP.

PI Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T;

DR WPI, 2004-711317/70.

DR N-PSDB; ADT9605.

XX New engineered alkaline protease with improved activity and thermal
PT stability, useful particularly in detergent such as laundry detergent.

XX Disclousre; Page 19-25; 35pp; English.

XX The invention relates to an alkaline protease and its encoding gene. The
CC alkaline protease can be expressed by standard recombinant methodology.
CC The alkaline protease is useful in the industry particularly in
CC detergents such as laundry detergent, but also in fiber modifying
CC agents, leather processing agents, cosmetic compositions, bath additives,
CC food-modifying agents, and pharmaceuticals. The enzyme has good activity
CC and thermal stability. The present sequence represents an alkaline
CC protease from *Bacillus* sp. KSM-KP43.

XX Sequence 641 AA;

Query Match 99.0%; Score 3029; DB 8; Length 641;

Best Local Similarity 93.3%; Pred. No. 1.6e-237;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 MRKKKKVFLSVLSAAAILSTVALKNPSAGAXXFDLDFKGIQTITDXXGFSKQGTGAAA 60
DB 1 MRKKKKVFLSVLSAAAILSTVALSNPSAGARNFDLDFKGIQTITDAGFSKQGTGAAA 60
QY 61 FLVSEENVKLGKGLKKLETPANNKXHXOFNGPILEETKQXLEXTGAKILDYIPDAY 120
DB 61 FLVSEENVKLPKGLQKLETPANNKXHXOFNGPILEETKQXLEXTGAKILDYIPDAY 120
QY 121 IVEYEGDVXSXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVKAALDTKQXKEVOLR 180
DB 121 IVEYEGDVKSANSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAALDTKQXKEVOLR 180
QY 181 GIEKIAQXXXSNDVYITTAKEPEYKVMNDVARGIVADVAOSSYGLYGGQGIIVAVADTGLD 240
DB 181 GIEKIAQFAISNDVYITTAKEPEYKVMNDVARGIVADVAOSSYGLYGGQGIIVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSSQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 QSIMDSGGGLGGLPSNLQTLFSSQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360
QY 361 TILFAAGNEXPNGGTITSAFGTAKNAITVGATENLPSFGSYADNINHYAQSFSRGPTDQ 420
DB 361 TILFAAGNEXPNGGTITSAFGTAKNAITVGATENLPSFGSYADNINHYAQSFSRGPTDQ 420
QY 421 RIKPVMAPGKTILSARSLAPDSSFWANHDSKAYMGSTMAFTIVAGNVAQLREHFVK 480
DB 421 RIKPVMAPGKTILSARSLAPDSSFWANHDSKAYMGSTMAFTIVAGNVAQLREHFVK 480
QY 481 NRGITPKPSLKLKALIAAGADXLGYPNGNQGCVTLTDKSLNVAVNESSLSTSQKAT 540
DB 481 NRGITPKPSLKLKALIAAGADXLGYPNGNQGCVTLTDKSLNVAVNESSLSTSQKAT 540
QY 541 YXFTATAGKPLKISLVMSDAPASTAATVTLVNDLVLVTAPENGIXYVANDFPKXXKXND 600
DB 541 YSFTATAGKPLKISLVMSDAPASTAATVTLVNDLVLVTAPENGIXYVANDFPKXXKXND 600
QY 601 GRNNVENVFINKPOSGTITIEVOAVNPVGPQESLAIYN 640
DB 601 GRNNVENVFINKPOSGTITIEVOAVNPVGPQESLAIYN 640

```

1

RESULT 5
AA17091
ID AA17091 standard: protein; 640 AA.

XX AA17091;
XX

DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)

XX Bacillus alkaline protease.

XX Alkaline protease; *Bacillus*; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent.

XX *Bacillus* sp.

XX W09918218-A1.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-JP004528.

XX 07-OCT-1997; 97JP-00274570.

XX (KAOS) KAO CORP.

PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;

DR WP1; 1999-287736/27.

DR N-PSDB; AAX37279.

PT Alkali protease from *Bacillus* used in washing powders.

PS Disclousre; Page 63-68; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of
CC *Bacillus*. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)

XX Sequence 640 AA;

Query Match 99.0%; Score 3028; DB 2; Length 640;

Best Local Similarity 93.3%; Pred. No. 1.9e-237;

Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 MRKKKYFLSVLSAAAILSTVALKNPSAGAXXFDLDFKGIQTITDXXGFSKQGTGAAA 60
DB 1 MRKKKYFLSVLSAAAILSTVALSNPSAGARNFDLDFKGIQTITDAGFSKQGTGAAA 60
QY 61 FLVSEENVKLGKGLKKLETPANNKXHXOFNGPILEETKQXLEXTGAKILDYIPDAY 120
DB 61 FLVSEENVKLPKGLQKLETPANNKXHXOFNGPILEETKQXLEXTGAKILDYIPDAY 120
QY 121 IVEYEGDVXSXXXXIEHVESVEPYLPYXXIDPOLFTKGASXLVKAALDTKQXKEVOLR 180
DB 121 IVEYEGDVKSATSTIEHVESVEPYLPYRIDPOLFTKGASXLVKAALDTKQXKEVOLR 180
QY 181 GIEKIAQXXXSNDVYITTAKEPEYKVMNDVARGIVADVAOSSYGLYGGQGIIVAVADTGLD 240
DB 181 GIEKIAQFAISNDVYITTAKEPEYKVMNDVARGIVADVAOSSYGLYGGQGIIVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALGRTNNANDTNGHGHVAGSVLGNSTNGKGAPOANLVF 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSSQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSSQASAGARIHTNSMGAAVNGAYTTDSRNVDYVRKNDM 360

```

```

Db      301 OSIMDSGGIGLPSNLQTLFSGQAYSAGARIHTNSWGAAVNGAYTTDSRVDDYVAKNDM 360
QY      361 TLIFAGNEXPNPGTISAPGTAKNAITVGA TENLRPSGSYADINIHVAQFSSRGPTKDG 420
Db      361 TLIFAGNEXPNPGTISAPGTAKNAITVGA TENLRPSGSYADINIHVAQFSSRGPTKDG 420
QY      421 RIKPDVMAFGTILISARSSLPDSSFWANHDSKYAYMGTSMA TPVAGVNAQLREHFYK 480
Db      421 RIKPDVMAFGTILISARSSLPDSSFWANHDSKYAYMGTSMA TPVAGVNAQLREHFYK 480
QY      481 NNGITPKPSILKKAALVAGAADIGLGPNGQGWGRVTLDKSLNVAAYNESSLSTSQKAT 540
Db      481 NNGITPKPSILKKAALVAGAADIGLGPNGQGWGRVTLDKSLNVAAYNESSLSTSQKAT 540
QY      541 YKFTATAGKPLKISLWSDAPASTASTVTLVNDLDTVITAPNGTXYVGDPEKPKXXKMD 600
Db      541 YKFTATAGKPLKISLWSDAPASTASTVTLVNDLDTVITAPNGTXYVGDPEKPKXXKMD 600
QY      601 GRNNVENVFINKPQSGTYTIEVOAYNVVPGQKFSIAIYN 640
Db      601 GRNNVENVFINKPQSGTYTIEVOAYNVVPGQKFSIAIYN 640

RESULT 6
AA117089
ID      AA117089 standard; protein; 639 AA.
XX
AC      AA117089;
XX
DT      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX
DE      Bacillus alkaline protease.
KW      Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW      washing composition; oxidising agent.
OS      Bacillus sp.
PN      WO9118218-A1.
PD      15-APR-1999.
PF      07-OCT-1998; 98WO-JP004528.
PR      07-OCT-1997; 97JP-00274570.
XX
(PAOS ) KAO CORP.
XX
PA      Takaiwa M, Okuda M, Saeki K, Kubota H, Hironi J, Kageyama Y;
PI      Shikata S, Nomura M;
DR      WPI, 1999-287736/27.
DR      N-PSDB; AAX37277.
XX
PT      Alkali protease from Bacillus used in washing powders.
XX
PS      Disclosure; Page 53-58; 71pp; Japanese.
XX
CC      The invention relates to alkaline proteases produced by strains of
CC      Bacillus. The proteases ability to digest casein is not inhibited by
CC      oleic acid and they have a high stability to oxidising agents. The
CC      alkaline protease of the invention has the following properties: (a) it
CC      is active over the pH range 4-13 and has at least 80% of its optimum
CC      activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC      stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC      its ability to digest casein is not inhibited by oleic acid; (e) it has
CC      molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC      used as enzymes in washing compositions for use in automatic dishwashers
CC      and for washing clothes. The stability to oxidising agents allows the
CC      enzyme to be an effective component of washing compositions including
CC      bleaches. The present sequence represents an alkaline protease. (Updated
CC      on 20-MAR-2003 to correct DR field.)

```

```

XX      XX
SQ      Sequence 639 AA;
Query Match      98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 8.7e-237;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY      4 KKKVFLSVLSAAALISTVALKNPSAGXAXFPLDPFGIGTITDDXXGFSKXOTGAAPLV 63
Db      3 KKKVFLSVLSAAALISTVALKNPSAGDARTPLDFKIGITTDVSGFSKOTGAAPLV 62
QY      64 ESENVTLKKGLKKKLETPANNKLIHXFNGPILBETKOLXETGAKLIDYIPDYAYVE 123
Db      63 ESENVTLKKGLKKKLETPANNKLIHXFNGPILBETKOLXETGAKLIDYIPDYAYVE 122
QY      124 YEGDYKSKXXXIEHVESVPEVLPYXXIDPQLFTKASXUVKAXALDTOKXNKEVOLRGIE 183
Db      123 YEGDVQSKVRSIEHVESVPEVLPYKYIDPQLFTKASLTVKALALDTOKNKEVOLRGIE 182
QY      184 XIAQXXXSNDVYVITAKPEYKYMNDVARGIVKADVAQSSYGLYGQQLVAVADTGLDTR 243
Db      183 EIAQVVASNDVYVITAKPEYKYMNDVARGIVKADVAQSSYGLYGQQLVAVADTGLDTR 242
QY      244 NDSMHEAFRGKITLALYALGRTNNANDTNGSTHVASGLGNGXTKGMAPQANLVFQSI 303
Db      243 NDSMHEAFRGKITLALYALGRTNNANDTNGSTHVASGLGNGXTKGMAPQANLVFQSI 302
QY      304 MDSXGGLGGLPSNLQTLFSGQAYSAGARIHTNSWGAAVNGAYTTDSRVDDYVAKNDMTIL 363
Db      303 MDSXGGLGGLPSNLQTLFSGQAYSAGARIHTNSWGAAVNGAYTTDSRVDDYVAKNDMTIL 362
QY      364 PAAGNEXPNPGTISAPGTAKNAITVGA TENLRPSGSYADINIHVAQFSSRGPTKDGRIK 423
Db      363 PAAGNEXPNPGTISAPGTAKNAITVGA TENLRPSGSYADINIHVAQFSSRGPTKDGRIK 422
QY      424 PVVMAFGTYIISARSSLPDSSFWANHDSKYAYMGTSMA TPVAGVNAQLREHFYKRG 483
Db      423 PVVMAFGTYIISARSSLPDSSFWANHDSKYAYMGTSMA TPVAGVNAQLREHFYKRG 482
QY      484 ITPKPSILKKAALVAGAADXGLGYPNGQGWGRVTLDKSLNVAAYNESSLSTSQKATYXF 543
Db      483 ITPKPSILKKAALVAGAADVGLGYPNGQGWGRVTLDKSLNVAAYNESSLSTSQKATYXF 542
QY      544 TATAGKPLKISLWSDAPASTASTVTLVNDLDTVITAPNGTXYVGDPEKPKXXKMDGN 603
Db      543 TATAGKPLKISLWSDAPASTASTVTLVNDLDTVITAPNGTXYVGDPEKPKXXKMDGN 602
QY      604 NVENVFINKPQSGTYTIEVOAYNVVPGQKFSIAIYN 640
Db      603 NVENVFINKPQSGTYTIEVOAYNVVPGQKFSIAIYN 639

RESULT 7
AA117088
ID      AA117088 standard; protein; 640 AA.
XX
AC      AA117088;
XX
DT      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX
DE      An alkaline protease sequence from Bacillus species.
KW      Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW      washing composition; oxidising agent.
OS      Bacillus sp.
XX
FH      Key      Location/Qualifiers
FT      Misc-difference 1..640
FT      note="all residues indicated as Xaa are arbitrary amino
FT      acids"
XX

```

```

PN WO9918218-A1.
XX 15-APR-1999.
XX 07-OCT-1998; 98WO-JP004528.
XX 07-OCT-1997; 97JP-00274570.
XX (KAOS ) KAO CORP.
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX Shikata S, Nomura M;
XX WPI: 1999-287736/27.
XX N-PSDB; AAX37278.
XX Alkali protease from Bacillus used in washing powders.
XX Claim 3; Page 50-53; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;

Query Match          98.6%; Score 3016; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.8e-236;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRXKKKVFSLVLSAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXQJTGAAA 60
DB 1 MRXKKKVFSLVLSAAILSTVALXNPSAGXARXFDLDFKGIQTITDXXGFSKXQJTGAAA 60
QY 61 FLVESENVTLKXGKLGKLETVPANNTKLTIXQFNGPILEETKQXLEXTGAKIIDYIPDYAX 120
DB 61 FLVESENVTLKXGKLGKLETVPANNTKLTIXQFNGPILEETKQXLEXTGAKIIDYIPDYAX 120
QY 121 IVEYEGDVYKXXXXIEHVESVEPYL.PXYXIDPOLFTKGASXIVKAXALDTKQXNREVOLR 180
DB 121 IVEYEGDVYKXXXXIEHVESVEPYL.PXYXIDPOLFTKGASXIVKAXALDTKQXNREVOLR 180
QY 181 GIEIXIAQXXXXSNDVXYITAKPEYKVMNDVARGIVADVAQSSYGLYGQGOIYAVADTGLD 240
DB 181 GIEIXIAQXXXXSNDVXYITAKPEYKVMNDVARGIVADVAQSSYGLYGQGOIYAVADTGLD 240
QY 241 TGRNDSMHEARFRGITLALYALGRTNMNDTNGHGHVAGSTLGNKTKMGAPOANLVE 300
DB 241 TGRNDSMHEARFRGITLALYALGRTNMNDTNGHGHVAGSTLGNKTKMGAPOANLVE 300
QY 301 OSIMDSXGGLGGLPNTLQTLFQOAXSAGARITNNSMGAAVANGAYTTDSRNVDYVRKNDM 360
DB 301 OSIMDSXGGLGGLPNTLQTLFQOAXSAGARITNNSMGAAVANGAYTTDSRNVDYVRKNDM 360
QY 361 TILFAAGNEXPMGRTISAPGTAKNAITVGATENLPPSFSGYADNINHVAFSSRPTXDG 420
DB 361 TILFAAGNEXPMGRTISAPGTAKNAITVGATENLPPSFSGYADNINHVAFSSRPTXDG 420
QY 421 RIKPVMAPGTIILARSILADSSFMANHDSKIVAMGTSMAATPIVAGNVAQLREHPTK 480
DB 421 RIKPVMAPGTIILARSILADSSFMANHDSKIVAMGTSMAATPIVAGNVAQLREHPTK 480

```

```

QY 481 NRGITPKPSLILKALILAGAADXGLGYPNGNGMGWRVTLDDKSLNAYVNESSXLSTSOXAT 540
DB 481 NRGITPKPSLILKALILAGAADXGLGYPNGNGMGWRVTLDDKSLNAYVNESSXLSTSOXAT 540
QY 541 YXFTATAGKPKIKSLVMSDAPASTTASVTLVNDLIVITAPNGTXYYGNDPFXXPXXKND 600
DB 541 YXFTATAGKPKIKSLVMSDAPASTTASVTLVNDLIVITAPNGTXYYGNDPFXXPXXKND 600
QY 601 GRNNVENVFINKPQSGTYTIEVOAIVNPVPGQXESLAIYN 640
DB 601 GRNNVENVFINKPQSGTYTIEVOAIVNPVPGQXESLAIYN 640

RESULT 8
AA17087
ID AA17087 standard; proteain. 639 AA.
XX
XX AA17087;
XX
XX 20-MAR-2003 (revised)
XX 21-JUL-1999 (first entry)
XX
DE An alkaline protease sequence from Bacillus species.
XX
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX washing composition; oxidising agent.
XX
OS Bacillus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..639
FT /note= "all residues indicated as Xaa are arbitrary amino
FT acids"
XX
XX WO9918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX Shikata S, Nomura M;
XX WPI: 1999-287736/27.
XX N-PSDB; AAX37277.
XX
XX Alkali protease from Bacillus used in washing powders.
XX Claim 3; Page 47-50; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
XX Bacillus. The proteases ability to digest casein is not inhibited by
XX oleic acid and they have a high stability to oxidising agents. The
XX alkaline protease of the invention has the following properties: (a) it
XX is active over the pH range 4-13 and has at least 80% of its optimum
XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
XX stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
XX its ability to digest casein is not inhibited by oleic acid; (e) it has
XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX used as enzymes in washing compositions for use in automatic dishwashers
XX and for washing clothes. The stability to oxidising agents allows the
XX enzyme to be an effective component of washing compositions including
XX bleaches. The present sequence represents an alkaline protease of the
XX invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 639 AA;

Query Match          98.3%; Score 3007; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 9.9e-236;

```


CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090

XX
 XX
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVAGIVKADVAGSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 266
 DB 1 NDVAGIVKADVAGSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 60
 QY 267 NANDTNGHGTTHVAGSVLGNKTNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSGQXS 326
 DB 61 NANDTNGHGTTHVAGSVLGNKTNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSGQXS 120
 QY 327 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAAGNEXPNGGTISAPGTAKNAI 386
 DB 121 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAAGNEXPNGGTISAPGTAKNAI 180
 QY 387 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSLAPDSSF 446
 DB 181 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSLAPDSSF 240
 QY 447 WANHDSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALIALAGADVGLGY 506
 DB 241 WANHDSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALIALAGADVGLGY 300
 QY 507 PNGNGMGRVTLDKSLNVAAYNESXLSQKATYXFTATAGKPLKISLWSDAPASTTA 566
 DB 301 PNGNGMGRVTLDKSLNVAAYNESXLSQKATYXFTATAGKPLKISLWSDAPASTTA 360
 QY 567 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVQAYN 626
 DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVQAYN 420
 QY 627 VPVGQXFSLAIVN 640
 DB 421 VPVGQXFSLAIVN 434

RESULT 12
 AAM50081

ID AAM50081 standard; protein; 434 AA.

XX
 AC AAM50081;

XX 12-AUG-2002 (first entry)

XX DE *Bacillus* sp KSM-KP9860 alkaline protease protein fragment.

XX KM Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX OS *Bacillus* sp.

XX PN EP1209233-A2.

XX PD 29-MAY-2002.

XX PF 22-NOV-2001; 2001EP-00127851.

XX PR 22-NOV-2000; 2000JP-00355156.

XX PR 12-APR-2001; 2001JP-00114048.

XX PA (KAO) KAO CORP.

XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Arai H, Sumitomo N;

PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.

XX PT New modified alkaline proteases useful in detergent compositions.

XX PS Claim 5; Page 12-13; 25pp; English.

XX
 CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC *Bacillus* sp strain KSM-KP9860 described in the method of the invention

XX
 SQ Sequence 434 AA;

Query Match 70.4%; Score 2155; DB 5; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVAGIVKADVAGSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 266
 DB 1 NDVAGIVKADVAGSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITLALYALGRN 60
 QY 267 NANDTNGHGTTHVAGSVLGNKTNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSGQXS 326
 DB 61 NANDTNGHGTTHVAGSVLGNKTNKGMAPQANLVFQSIIMDSXGGLGGLPSNLQTLFSGQXS 120
 QY 327 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAAGNEXPNGGTISAPGTAKNAI 386
 DB 121 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFPAAGNEXPNGGTISAPGTAKNAI 180
 QY 387 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSLAPDSSF 446
 DB 181 TVGATENLRPSFGSYADININHAQSSSRGPTDGRIKPDVMAPGTYIISARSLAPDSSF 240
 QY 447 WANHDSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALIALAGADVGLGY 506
 DB 241 WANHDSKAYMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALIALAGADVGLGY 300
 QY 507 PNGNGMGRVTLDKSLNVAAYNESXLSQKATYXFTATAGKPLKISLWSDAPASTTA 566
 DB 301 PNGNGMGRVTLDKSLNVAAYNESXLSQKATYXFTATAGKPLKISLWSDAPASTTA 360
 QY 567 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVQAYN 626
 DB 361 SVTLVNDLDLVITAPNGTXYVGNDFXKPKXKXWDRNNVENVFINKPQSGTYTIEVQAYN 420
 QY 627 VPVGQXFSLAIVN 640
 DB 421 VPVGQXFSLAIVN 434

RESULT 13

ID ADY33778 standard; protein; 434 AA.

XX AC ADY33778;

XX DT 05-MAY-2005 (first entry)

XX DE *Bacillus* species alkaline protease.

XX KM mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 XX bleaching detergent.

XX OS *Bacillus* sp.

XX PN EP1347044-A2.

XX PD 24-SEP-2003.

XX

PF 21-MAR-2003; 2003EP-00006472.
 XX
 XX 22-MAR-2002; 2002JP-00081428.
 PR 06-JUN-2002; 2002JP-00165987.
 PR 18-OCT-2002; 2002JP-00304230.
 PR 18-OCT-2002; 2002JP-00304231.
 XX
 PA (KAOS) KAO CORP.
 PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
 PI Kobayashi T, Nomura M;
 XX
 DR WPI; 2003-846540/79.
 DR N-PSDB; ADY33779.
 XX
 PT New alkaline protease having specific amino acid residue at a specific
 PT position of its amino acid sequence, useful for producing detergent
 PT compositions, laundry detergent, fiber modifiers, leather-treating agents
 PT or pipe cleaners.
 XX
 PS Claim 1; SEQ ID NO 1; 31pp; English.
 XX
 CC The invention relates to an alkaline protease having a fully defined
 CC sequence of 434 amino acids (I) given in the specification, or an amino
 CC acid sequence at least 80% homology with (I), where an amino acid residue
 CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
 CC selected from 46 amino acid residues. The amino acid residues at the
 CC corresponding positions are selected from: position 65: proline; position
 CC 101: asparagine; position 163: histidine, aspartic acid, phenylalanine,
 CC lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or
 CC valine; position 170: valine or leucine; position 171: alanine, glutamic
 CC acid, glycine or threonine; position 273: isoleucine, glycine or
 CC threonine; position 320: phenylalanine, valine, threonine, leucine,
 CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
 CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
 CC acid, arginine or histidine. The alkaline protease is useful for the
 CC production of a detergent composition, such as laundry detergent, fiber
 CC modifiers, leather-treating agents, cosmetic compositions, bath
 CC additives, food modifiers and pharmaceutical compositions. The alkaline
 CC protease may also be used as bleaching detergent, hard surface cleaning
 CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
 CC cleanser for medical tools. The new alkaline protease has a more potent
 CC proteolytic capacity, exhibiting excellent detergency for the removal of
 CC a complex stain, and has high secretion capacity. This sequence
 CC corresponds to the *Bacillus* sp. alkaline protease.
 CC
 XX
 SQ Sequence 434 AA;
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 207 NDVARGIVKADVAOSSYGLYGGQGVAVADTGLDTRGNDSSMHEAFRGIKITALYALGRTN 266
 DB 1 NDVARGIVKADVAOSSYGLYGGQGVAVADTGLDTRGNDSSMHEAFRGIKITALYALGRTN 60
 QY 267 NANTNGHGTHTVAGSYLNGXTNKGMAPANVVFOSIMDSKGLGSPENLQTLFSQAXS 326
 DB 61 NANTNGHGTHTVAGSYLNGXTNKGMAPANVVFOSIMDSKGLGSPENLQTLFSQAXS 120
 QY 327 AGARLHTNSMGAANVGAAYTTDSRNVDDYRKADMTLLFAAGNEXPFGITSPATAKAI 386
 DB 121 AGARLHTNSMGAANVGAAYTTDSRNVDDYRKADMTLLFAAGNEXPFGITSPATAKAI 180
 QY 387 TVGATPTNLRPFSGSYADNINHVAFSSSRGPTKGRIKPDVMAFGTYILSARSLAPDSGF 446
 DB 121 TVGATPTNLRPFSGSYADNINHVAFSSSRGPTKGRIKPDVMAFGTYILSARSLAPDSGF 240
 QY 447 WANHDSKTYAMGISTMATIYVAGNVAQLREHFVYKRGITPKSLLKALIALGAAXGLCY 506
 DB 241 WANHDSKTYAMGISTMATIYVAGNVAQLREHFVYKRGITPKSLLKALIALGAAXGLCY 300
 QY 507 PNGNGMGKRVTLDSKLVNVAVNESSXLSTSQATYXFTATACKPLKSLVMSDAPASTTA 566

DB 301 PNEGNGMRVTLDSKLVNVAVNESSXLSTSQATYXFTATACKPLKSLVMSDAPASTTA 360
 QY 567 SVTLVNDLDTVTAPNGTXVYVGNDFKXXPKXNMDGRNVENVFINXPGSGTYTIEVQAYN 626
 DB 361 SVTLVNDLDTVTAPNGTXVYVGNDFKXXPKXNMDGRNVENVFINXPGSGTYTIEVQAYN 420
 QY 627 VPVGPQAFSLAIYN 640
 DB 421 VPVGPQAFSLAIYN 434
 RESULT 14
 ADZ51758
 ID ADZ51758 standard; protein; 434 AA.
 XX
 AC ADZ51758;
 XX
 DT 16-JUN-2005 (first entry)
 XX
 DE Mutant *Bacillus* sp. alkali protease, SEQ ID No:2.
 XX
 KM alkali protease; enzyme; surfactant; mutagenesis; mutein.
 XX
 OS *Bacillus* sp.
 OS Synthetic.
 XX
 PN JP2003125783-A.
 XX
 PD 07-MAY-2003.
 XX
 PF 26-OCT-2001; 2001JP-00329472.
 XX
 PR 26-OCT-2001; 2001JP-00329472.
 XX
 PA (KAOS) KAO CORP.
 XX
 DR WPI; 2003-856569/80.
 XX
 PT New alkali protease useful for preparing detergents comprises
 PT substitution mutations at 251 or 256 position of protease KP43 derived
 PT from *Bacillus* species KSM-KP43.
 XX
 PS Claim 3; SEQ ID NO 2; 16pp; Japanese.
 XX
 CC The invention relates to a mutant alkali protease having an amino acid
 CC deletion at position(s) 251 or 256 in a fully defined sequence given as
 CC SEQ ID No:1 in the specification, or the following amino acid residue
 CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
 CC amino acid residue Lys, Ser, Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
 CC Glu at position 256. Also described are: (i) an alkali protease having to
 CC the amino acid sequence of SEQ ID No:1, or having 60% or more homology to
 CC SEQ ID No:1, or a deletion of amino acid corresponding to position 251 or
 CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
 CC vector containing the gene, and (iv) a transformed organism containing
 CC the recombinant vector. The alkali protease sequence having 60% or more
 CC homology to SEQ ID No:1 is chosen from a fully defined sequence selected
 CC from SEQ ID Nos 2-7 as given in the specification. The transformed
 CC organism is a microorganism. The mutant alkali protease is useful for
 CC preparing detergents. The mutant alkali protease exhibits high resistance
 CC against oxidizing agent. The mutant alkali protease has high specific
 CC activity. This sequence represents a mutant *Bacillus* sp. alkali protease.
 XX
 SQ Sequence 434 AA;
 Query Match 70.4%; Score 2155; DB 7; Length 434;
 Best Local Similarity 96.3%; Pred. No. 1.3e-166;
 Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 QY 207 NDVARGIVKADVAOSSYGLYGGQGVAVADTGLDTRGNDSSMHEAFRGIKITALYALGRTN 266
 DB 1 NDVARGIVKADVAOSSYGLYGGQGVAVADTGLDTRGNDSSMHEAFRGIKITALYALGRTN 60


```

QY 267 NNDNTNGHGTAVAGSVLGNSTKNGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 326
DB 61 NNDNTNGHGTAVAGSVLGNSTKNGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 120
QY 327 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLPAAGNEXPNGGTTISAPGTAKNAI 386
DB 121 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLPAAGNEXPNGGTTISAPGTAKNAI 180
QY 387 TVGATENTLRPSFGSYADNINHVAQPSRRPTKGRIKEDVMAFGTYILSARSSLAPDSSF 446
DB 181 TVGATENTLRPSFGSYADNINHVAQPSRRPTKGRIKEDVMAFGTYILSARSSLAPDSSF 240
QY 447 WANHSKAYVMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAAADIIGLGY 506
DB 241 WANHSKAYVMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAAADIIGLGY 300
QY 507 PNGNGQWGRVTLDKSLNVAAYNNESSXLTSTOKATYXFTATAGKPKLISLWSDAPASTTA 566
DB 301 PNGNGQWGRVTLDKSLNVAAYNNESSXLTSTOKATYXFTATAGKPKLISLWSDAPASTTA 360
QY 567 SFTLVNDLVLVTAPNGTYVGNDFPKXPKXNMGRNNVNFINKPQSGTTTIEVOAYN 626
DB 361 SFTLVNDLVLVTAPNGTYVGNDFPKXPKXNMGRNNVNFINKPQSGTTTIEVOAYN 420
QY 627 VPVGPOXPSLAIVN 640
DB 421 VPVGPOXPSLAIVN 434

RESULT 15
ADZ51757
ID ADZ51757 standard; protein; 434 AA.
AC ADZ51757;
DT 16-JUN-2005 (first entry)
XX
DE Wild-type Bacillus sp. alkali protease.
KM alkali protease; enzyme; surfactant; mutagenesis.
XX
OS Bacillus sp.
XX
PN JP2003125783-A.
XX
PD 07-MAY-2003.
XX
PF 26-OCT-2001; 2001JP-00329472.
XX
PR 26-OCT-2001; 2001JP-00329472.
XX
PA (KAOS ) KAO CORP.
XX
DR WPI; 2003-856569/80.
XX
PT New alkali protease useful for preparing detergents comprises
PT substitution mutations at 251 or 256 position of protease KP43 derived
PT from Bacillus species KSM-KP43.
XX
PS Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
CC The invention relates to a mutant alkali protease having an amino acid
CC deletion at position(s) 251 or 256 in a fully defined sequence given as
CC SEQ ID No:1 in the specification, or the following amino acid residue
CC Asp, Gly, Ala, Asp, Thr, Ile, Val, Leu or Glu at position 251, or the
CC amino acid residue Lys, Ser, Phe, Arg, Tyr, Met, Thr, Ile, Val, Leu, or
CC Glu at position 256. Also described are: (i) an alkali protease having
CC the amino acid sequence of SEQ ID No:1, or having 60% or more homology to
CC SEQ ID No:1, or a deletion of amino acid corresponding to position 251 or
CC 256, (ii) a gene which encodes an alkali protease, (iii) a recombinant
CC vector containing the gene, and (iv) a transformed organism containing
CC the recombinant vector. The alkali protease sequence having 60% or more
CC homology to SEQ ID No:1 is chosen from a fully defined sequence selected

```

```

CC from SEQ ID Nos 2-7 as given in the specification. The transformed
CC organism is a microorganism. The mutant alkali protease is useful for
CC preparing detergents. The mutant alkali protease exhibits high resistance
CC against oxidizing agent. The mutant alkali protease has high specific
CC activity. This sequence represents wild-type Bacillus sp. alkali
CC protease.
XX
SQ Sequence 434 AA:
Query Match 70.4%; Score 2155; DB 7; Length 434;
Best Local Similarity 96.3%; Pred. No. 1,3e-166;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 207 NDVARGIVADVAQSSYGLYGQGIIVAAVDTGLDGRNDSMHBAFRGKITLVALGRTN 266
DB 1 NDVARGIVADVAQSSYGLYGQGIIVAAVDTGLDGRNDSMHBAFRGKITLVALGRTN 60
QY 267 NNDNTNGHGTAVAGSVLGNSTKNGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 326
DB 61 NNDNTNGHGTAVAGSVLGNSTKNGMAPOANLVFOSIMDSXGGLGSLPSNLQTLTLPFOAYS 120
QY 327 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLPAAGNEXPNGGTTISAPGTAKNAI 386
DB 121 AGARLHTNSWGAAVNGAYTTDSRNVDYVRKNDMTLLPAAGNEXPNGGTTISAPGTAKNAI 180
QY 387 TVGATENTLRPSFGSYADNINHVAQPSRRPTKGRIKEDVMAFGTYILSARSSLAPDSSF 446
DB 181 TVGATENTLRPSFGSYADNINHVAQPSRRPTKGRIKEDVMAFGTYILSARSSLAPDSSF 240
QY 447 WANHSKAYVMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAAADIIGLGY 506
DB 241 WANHSKAYVMGGSMTATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAAADIIGLGY 300
QY 507 PNGNGQWGRVTLDKSLNVAAYNNESSXLTSTOKATYXFTATAGKPKLISLWSDAPASTTA 566
DB 301 PNGNGQWGRVTLDKSLNVAAYNNESSXLTSTOKATYXFTATAGKPKLISLWSDAPASTTA 360
QY 567 SFTLVNDLVLVTAPNGTYVGNDFPKXPKXNMGRNNVNFINKPQSGTTTIEVOAYN 626
DB 361 SFTLVNDLVLVTAPNGTYVGNDFPKXPKXNMGRNNVNFINKPQSGTTTIEVOAYN 420
QY 627 VPVGPOXPSLAIVN 640
DB 421 VPVGPOXPSLAIVN 434

```

Search completed: April 7, 2006, 17:37:31
Job time : 139.608 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:37:51 ; Search time 27.5215 Seconds
(without alignments)
2237.477 Million cell updates/sec

Title: US-10-784-870-2

Perfect score: 3059

Sequence: 1 MRXKKKVLSTLSAAALST.....RYQAVNPVGPQXPSLATVN 640

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Libbing filter 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518.5	15.9	1743	2 T18279	multidrug resistan
2	477.5	15.6	1905	2 T18267	multidrug resistan
3	350.5	11.5	444	2 B83891	intracellular alka
4	319.5	10.4	442	2 A69587	pyrolysins (EC 3.4.
5	297.5	9.7	1398	2 T28159	microbial serine p
6	286	9.3	806	2 A41341	subtilisin (EC 3.4
7	279.5	9.1	419	1 S25835	subtilisin-type pr
8	276.5	9.0	799	2 G83753	subtilisin (EC 3.4
9	274.5	9.0	420	1 S23407	serine proteinase
10	270	8.8	580	2 S11890	alkaline serine pr
11	268.5	8.8	715	2 J04908	cell wall-associat
12	262	8.6	894	2 P63730	surface layer-asso
13	260.5	8.5	1345	2 T29090	high-alkaline seri
14	257.5	8.4	378	2 A33973	aqualysin (EC 3.4.
15	255	8.3	513	1 A35742	high-alkaline seri
16	250.5	8.2	380	2 A49778	serine proteinase
17	250.5	8.2	601	2 J04576	subtilisin (EC 3.4
18	250	8.2	382	2 J39780	subtilisin (EC 3.4
19	249.5	8.2	757	2 C84120	serine proteinase,
20	248	8.1	627	2 D75393	alkaline proteinase
21	246.5	8.1	402	1 J00332	subtilisin (EC 3.4
22	245.5	8.0	381	2 J00778	subtilisin (EC 3.4
23	245	8.0	382	1 SUBSN	subtilisin (EC 3.4
24	243.5	8.0	381	1 SUBSI	subtilisin (EC 3.4
25	242.5	7.9	379	1 SUBSCL	subtilisin (EC 3.4
26	242	7.9	1331	2 A72647	probable surface 1
27	241.5	7.9	381	1 SUBSS	subtilisin (EC 3.4
28	239.5	7.8	401	2 J39974	serine proteinase
29	239.5	7.8	534	1 J06173	alkaline proteinase

30	237.5	7.8	381	2 J01487	subtilisin (EC 3.4
31	235	7.7	519	2 S71451	haloalysin R4 (EC 3
32	232	7.6	488	2 A11930	protease (import
33	231.5	7.6	1167	1 A35066	serine proteinase
34	230	7.5	384	2 J04802	alkaline proteinase
35	223.5	7.3	409	1 S32905	serine proteinase
36	223	7.3	321	1 S27501	alkaline proteinase
37	219.5	7.2	613	2 S75976	hypothetical prote
38	218.5	7.1	272	2 A23624	subtilisin (EC 3.4
39	218.5	7.1	275	2 J01085	subtilisin (EC 3.4
40	218.5	7.1	533	1 J00146	serine proteinase
41	218.5	7.1	535	2 B83358	alkaline serine pr
42	218	7.1	361	2 G83756	subtilisin-type al
43	216.5	7.1	361	2 A48372	high-alkaline seri
44	216.5	7.1	525	2 G84406	haloalysin (importe
45	216.5	7.1	645	1 SUBSMP	serine proteinase

ALIGNMENTS

RESULT 1

multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Jul-2004
C:Accession: T18279
R:Shaulsky, G.; Loomis, W.F.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z18855
A:Accession: T18279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1743 <SHA>
A:Cross-references: UNIPROT:Q23868; UNIPARC:UP1000013687D; EMBL:U60086; NID:G1399914; P1
A:Gene: tagC

Query Match 16.9%; Score 518.5; DB 2; Length 1743;

Best Local Similarity 24.2%; Pred. No. 3.8e-28;
Matches 194; Conservative 114; Mismatches 213; Indels 281; Gaps 33;

QY	68	HIQFNGPIERTKQKL-----EXTGAKIIDYIPDYAYIV---EYEGDVXSK	131
DB	115	YIVQFKRINDTRQQLKPLIGTDIVLDEQYQSHIVHYIPHDSFLVMTQEQVLSLS	174
QY	132	XXXIHVESVEPY-----LPYXIDPOLFTKGASLYK--AXALDT--KQXN	174
DB	175	KEMVWIGEPFSPNNIHLNVEKSIQGLPVYII---LDSINSLQRMENLINSILKSYN	230
QY	175	KEVQLRGI-----EXIAQXXXSNDVXYITAKPEYK	204
DB	231	SKVKLTLLNQKQKSIYCNDESPSPCSLINSKLVYQWISQESN---FIERSEKQ	287
QY	205	VWMDVARGIV-----KADVAQSSYGLYGGQIVAVADPTGLDTR--NNS----	246
DB	288	TANRLSPKVFEGTKDTLVNNDKVDIP---LRGQILSLADTGLDGHCPSPDSKYP	342
QY	247	---SWHBAFGKITLVALGRTNANDNTNGHTYHAGSVLG-----NGXTNKGAPQA	296
DB	343	PLANSVNLNHR-KVVTYITTSDDSDKVDGHGTHICGSAAGPBEISSVNISFSLATDA	401
QY	297	NLVFQSIQSGKGLGL--PSNLQTLFQOAYSAGARITNSWGA-----AVNGAYTTDSN	350
DB	402	KIAF---FDLASGSSSLTPPSDLKQLYQPLVDAGARVHCDSSGVSVEGYTSSYSDTAS	458
QY	351	VDDYVRKX-DWITILPAAGKXNGSTIS--APGTAKNITVCATENLR-----PSRFS	400
DB	459	IDDFLFTHPDFTILPAAGN---NEGYLSLTQSTAKNIVTGAHQTHENYITDGPNYIN	515
QY	401	YADNI-----	405
DB	516	YQSSVDINQELICDPDSRYKCYTTTAQCCLESNATTLGASCCPTLARKSVIDAANTOPILY	575

406 --NHVAQSSRPTKOGRIKPDVMAFGTYILSARSSLA-----PDSGFMANHDSKTYAM 457
Db 576 NENNISSPSSKGPTRDGRKRPALVAFGEYITSARNGANTTDQCCDGLS-PTYNLALA-I 633
Qy 458 GGTSMATPIVAGNVAQLREH-----FYKNGITPKPSELKRALIAGA----- 499
Db 634 SGTSMATPMAAATITLKOYLDVGYTPGSGIVESKLOPTGSLKALMINNAQLNGTFQ 693
Qy 500 --ADGGLGYPNGN-----QGWGRVTLDKSLNVAAYNESS----- 531
Db 694 LITSSITYPSPNQVFNPAASLVQGWGALIRMSNMLHYVNNNNNNNNKTSIDGITKPDGI 753
Qy 532 -----XISTSQKATYXFT-----ATNGKPKL--ISLWSDAPASTT 565
Db 754 GGLDLRLVYKPNQWKESSLSTGQNTSYCFYTKPSSSSSSNSGNIPRVVATLVTWTDPPSYAG 813
Qy 566 ASVTLVNDLVI-----TAPNGTYVGNDFXPPXXNMWDRNNVENVFINKP 613
Db 814 AKPNLVNLDLMTIYRRNGSTIFYSNQGSFGL-----LAPQDTLNNVEGIVHNPT 867
Qy 614 QSGTYTIEVOAYNVPVGPQXFS 635
Db 868 EPMTYRFVAVAGTNVMPGFQNF 889

RESULT 2
T18267
multidrug resistance protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18267
R:Shaulsky, G.; Kuspa, A.; Loomis, W.F.
submitted to the EMBL Data Library, January 1995
A:Description: An MDR transporter/serine protease gene is required for prestalk specia1
A:Reference number: Z18850
A:Accession: T18267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1905 <SHA>
A:Cross-references: UNIPROT:P54683; UNIPARC:UPI000013687C; EMBL:U20432; NID:g664839; PID
C:Genetics:
A:Gene: tagB

Query Match 15.6%; Score 477.5; DB 2; Length 1905;
Best Local Similarity 24.5%; Pred. No. 3.2e-25;
Matches 193; Conservative 98; Mismatches 237; Indels 259; Gaps 31;
88 HIXQFNPIRETKQXLE---XTGAKILDYIPDYAVIVEX-----EGDYXS 130
Db 172 YIVQFKRINDREQLKEPLIGTDITLKEOPFKSHIVHYIPHDSFLVMTKQSVLLS 231
Qy 131 XXXXIEHVESVEPY-----LPXYXI-----DPOLFTKGSXLT-----V 163
Db 232 SKEMISWIGHEPSPKIHNLNHEKSIQVPIYIILSGTNSLIQRMENLNLISVNSKV 291
Qy 164 KAXALDTQ-----XNKEVQLGEIEXIAQXXKS--NDVXYITTAPEYKVM 206
Db 292 KLTILINQKLSIYCNDESPSSSSSSCSLIGSEKIYKWIISQESSENYIERSEKQTA 351
Qy 207 NDVARGIV-----KADVAQSSYGLGGQGIYVADTGIDTGR-----NDS----- 246
Db 352 NRLEPTVYFGTKDKLVNNDRIDP-----LNGKGQILSIADTGLGSHCFPSDSKTYPIPF 406
Qy 247 -SMHEAFRGKITALYALGRTNNANDTNGHGTIVAGSVLNGXTN-----KGMAPQANL 298
Db 407 NQVNEHNRKVT---YIYHNEEDYVNGHGVGSAAGTPEDSSMAISPSGLATAKI 463
Qy 299 VFGSINDXGGLGGLPSNLQTLFSGQASAGARHTNSWGA---AVNGAYTTDSRNVDDY 354
Db 464 AFYD-LSSGSSEPTPEPDYSQMYKPLYDAGARVHGDWSGVSLOQYGYGSDDAGGIDAF 522
Qy 355 VRK-NDMTILFAAGNEKXNGGTISAPGAKNAITVGAENLRPSGS-----YADNI- 405

Db 523 LYEYPERSLPAAGNN-ELFASLILQATAKNAITVGAEQTAHVNVSDALRYDESDAN 581
Qy 406 -----NHVAQF 411
Db 582 FQRPCLPDKKVCNTTAKCGSESVNKGDLCCPASIKNQNSDFTTQPPQYNNMNSGF 641
Qy 412 SSRGPTKDGRIKPDVMAFGTYILSARSS-----LAPDSGFMANHDSKTYAMGTSMATP 465
Db 642 SSKGPTHDGRKLPDIVAFGEYITSARNGENSTDOCGGSL--FNANGLMSISGSMATP 699
Qy 466 IYAGNVAQLREH-----YKNGITPKPSELKRALI----- 496
Db 700 LATATITLKOYLDVGFPTGESVVENKLTFTGSLKALMINNAQLNGTYFWSNASSTNP 759
Qy 497 AGAADXGLGYPNGQWGRVTLDKSLNVAAYNESS-----XLISTSQKAT----- 540
Db 760 SNAIFRQINGANLLOQMGLAMN---NMLYKSSNPTRPSRWIGIGLGRKQKATWEMED 816
Qy 541 -----YXFT-----ATNGKP-LKISLWSDAPASTTASVTLVNDLVI-- 576
Db 817 SLSSGLNKSXCFYTKPSSSSSGSGGGTFRIVATLVTWTDPPSYGAKFNLVNLDDL 876
Qy 577 -----VTTAPN--GTXYVGNDFXPPXXNMWDRNNVENVFINKPQSGTYTIEVOAYNVP 628
Db 877 NSDDSIITIGNSGSLQPAKTAQ-----DTLNNVEGIIITPKANNYFTIAGTNVP 931
Qy 629 VGPQXFS 635
Db 932 IGPKFS 938

RESULT 3

883891
intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain C

C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: B83891
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: UNIPROT:Q9KBJ7; UNIPARC:UPI000003D43; GB:AP001513; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: aprX

Query Match 11.5%; Score 350.5; DB 2; Length 444;
Best Local Similarity 30.2%; Pred. No. 3.9e-17;
Matches 114; Conservative 54; Mismatches 127; Indels 83; Gaps 17;
183 EXIAQXXKSDVXYITTKPEYKMNDAKGIYVA-DVAQSSYGLYGGQIYVADTGIDT 241
Db 100 ESIQEMLVCKDIRKITYLNREYHALDTPAVESAQAFVIRNGETLTGQDVIAVIDGT-- 157
Qy 242 GRNDSMHEAFRGKITALY-ALGRTNNANDTNGHGTIVAGSVLNGXTN-----KGMAPQ 296
Db 158 -----YHEDLEGRIKAFVDFVNRREPYDDNGHGTICAGDMAAGNASSDQGTGRPAEA 212
Qy 297 NLVFGSINDXGGLGGLPSNLQTLFSGQASAGARHTNSWGAAYNVAAYTTDSRNV----- 351
Db 213 NVIGVATLNRQ--GWSGLESLIMGV-----EMCIQYNEBHPDDPIHIIISML 257
Qy 352 -----DDYVF-----KNDTILFAAGNEKXNGGTISAPGTAKNAITVGAATEN 393
Db 258 GQALPYENBOEPDPMVAVIVEANWAGITVCAAGNSGPDQQTASPSGVSEKRVITVGLDD 317
Qy 394 LRPFSGYADNINVAQSSRGPTKDGRIKPDVMAFGTYILSARSSLAPOSSF-----WA 448

Db 318 -RDYTDREDD---VAPFSRGPTIYGKPKPDILAPGVNIISLRS---FNSFYDKIOKGS 370
QY 449 NNDKXAYMGSTSMATPIYAGNAQAREHFVKKRGITTPRSLKALALIGA---AD----- 501
Db 371 RVGSHYTMGSGTSMATPVCGVVALMLQ---EPNLTPEB-VKTRLMGSTRMADRDPN 425
QY 502 -XGLGY-----PENGQ 511
Db 426 VYGAGYISABGAIIPNSEE 443

RESULT 4

intracellular alkaline serine proteinase aprx - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 09-Jul-2004
C/Accession: A69587
R/Kunitz, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beren
C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Ertling, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, R.; Roche, B.; Rose, M.; Sadele, Y.; Sato, T.; Scanlon,
A.; Schlegel, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tanakaishi, A.; Tanaka, T.; Teperstra, P.; Tognoni, A.; Tostato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumsel, B.; Yoshikawa, H.; Darchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; M01D:98044033; PMID:9384377
A/Accession: A69587
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-442 <KUN>
A/Cross-references: UNIPROT:O31788; UNIPARC:UPI000006043P; GB:Z99113; GB:AL009126; NID:9
A/Experimental source: strain 168
C/Genetics:
A/Gene: aprX
F/146-398/Domain: subtilisin homology <SBT>

Query Match 10.4%; Score 319.5; DB 2; Length 442;
Best Local Similarity 30.0%; Pred. No. 5.8e-15;
Matches 106; Conservative 46; Mismatches 108; Indels 93; Gaps 15;
QY 202 EYKAVNDVARGIVKA-DVAQSSYGLYGQGIYAVADDTGDRNDSMHEAFRGKITLY 260
Db 120 EYKALDRTTBSHAKVAVRANGCITLGKVTYAVVDGT-----YHPDLGR----- 168
QY 261 ALGRTNNAN-----DTNGHGHVAGSYLNGXTN-----KGNAPQANLVFQSIMSKGL 310
Db 169 -IGFADMTVQKTEPYDDNGHGHGTHCAGDVASGASSGQYRPAKRNALGVKVLNKGQ- 226
QY 311 GGLPBNLQTLFSGQASAGARIHTNSWGAAN-----GAYTTDSRNVDD 353
Db 227 GTLADITRGV-----EMCIQYNEDNDPEPIDIMSMISGGLALYDHEQEDP 272
QY 354 YVRKND-----MTLLFAGNEXPNGGTISAPGTAKNATTVGATENTLRFSGSYADNIH 407
Db 273 LVRAVEAKSAGIIVCVAGNAGSGPDSQITIASPVSERKVTYVGLADNDNTA-----SSDDT 328
QY 408 VAQFSRGPTQDGRIKPDVMAFGTILSARSSILAPDSSP-----WANHDSKYAYMGSTSM 462
Db 329 VASFSRSGPTVYGKPKPDILAPGVNIISLRS---FNSYIDKLOKSRVSGQFTMSGISM 385
QY 463 ATPYAGVNAQAREHFVKKRGITTPK--PSLLKALALIGAADYGLGYPNGNQW 513
Db 386 ATPICGIALALIQ---QNPDLTPDEVKELK-----NGTDMK 420

RESULT 5

T28159
pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C/Species: Pyrococcus furiosus
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004
C/Accession: T28159
R/Voorhorst, W.G.B.; Eggen, R.I.L.; Geertling, A.C.M.; Plateau, C.; Sizeren, R.J.; Vos,
J. Biol. Chem. 271, 20426-20431, 1996
A/Title: Isolation and characterization of the hyperthermostable serine protease, pyroly
A/Reference number: Z20481; M01D:9635370; PMID:8702780
A/Accession: T28159
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1398 <VOO>
A/Cross-references: UNIPROT:P72186; UNIPARC:UPI000016FDS9; EMBL:U55835; NID:G1556462; PI
A/Experimental source: DSM3638
C/Genetics:
A/Gene: D18
C/Keywords: hydrolase; serine proteinase

Query Match 9.7%; Score 297.5; DB 2; Length 1398;
Best Local Similarity 22.9%; Pred. No. 9.8e-13;
Matches 173; Conservative 76; Mismatches 248; Indels 257; Gaps 31;
QY 97 LERKQXLEXTGAKTLD-----YIPDYATVYR-----GDYKXXXXIXE 136
Db 77 LERATTELKLGABILDENRVLMVLKIKPKVKALNTSSLEKAMLRKYLSPPIVE 136
QY 137 -HVESVEPYLPYXIDPOLFT-----KGSALVKAYALDTKOKNKEV 177
Db 137 KQVTKKEP-----SLPKKYNSTWVINALQFIQEGYDQSGVAVVLDTGVDPNHFLSI 191
QY 178 QLRGLEXLAQKKXSDVAYITTKPEY-KYMN-----DYARGI----- 213
Db 192 FPDGRKKIIEWDFTEBGEVDTSFSPSKVNGTLINTPQVAGSLTNSGALMEYVK 251
QY 214 -----YKADVAGSSYGLY-----GOGQIYAVDT 237
Db 252 TVYVSNVTIGNTTSANGIYHFGLEPERFDLFPDQDQDFPVVLVNSGNGYDIAYVD 311
QY 238 GLDTRGNS-----SMHEAFRGKITLYALGRTNNAN-----DTNGHGHVAG 280
Db 312 DLDYDFTDEVPFGQYNTVYDAVPSYTYGPLMYVLAELIDPGEYAVPFGDGHGHVAG 371
QY 281 SYLNGXTN-----KGNAPQANLVFQSIMSKG 308
Db 372 TVAGDSNNDMWMLSMYSGEWVPSRLYGMDYTNVTTDYGVAFGAQIMAIRVLS-D 430
QY 309 GLGGLPSNLQTLFSGQASAGARIHTNSWGAANVGT--TDSRN--VDDYVRKNDMTILF 364
Db 431 GRGSMWDLIRGN-TYAAATHGADVLSMSLQG--NAPYLDGTDPESSAVDBLTERKGVVFT 487
QY 365 AAGNEXPNGGTISAPGTAKNATTVGATENTLRFSGSYAD-----NIN 406
Db 488 AAGNGPGINIVGSGVATKAITVGAAL-VPLINGVTVSQAALGYDYGFTYFPAYTNV- 545
QY 407 HVAQFSRGPTQDGRIKPDVMAFGTILSARSSILAPDSSPFWANHDSKYAYMGSTMATPI 466
Db 546 RIAPFSSRGPRIDGRIKENVAVPAGIYSSLPFWMGADF-----MSGSMATPH 595
QY 467 VAGNAQAREHFVKKRGITTPRSLKALALIGA-----DXGLGYPNGNGKGVNTIDK 520
Db 596 VSGVALLISG-PKEGIIYNPDIIKVLLEGATWLBGDPYTGQCTELDQHGVLNVTK 654
QY 521 SLNVAVYVNESKSLTSQKATYXFTTACKPLKISLWSDAPASTA---SVTLVNDLVL 577
Db 655 SMEI-----LAINGTTLPIYDHMAKYSIDPAEILGVDVIRGLYAR 696
QY 578 ITAPN-----GTYVGN-----DFXPXKXNN-----DG-----RNVE----- 606
Db 697 NSIPDIVEMHILKYVDTEYRTFEIYATBPWIKFPVSGSVILNNTBFLARVYKVEGLBP 756
QY 607 -----NVFINPKPGSTYTIIEVQANNVPVGPQXFS 635

Db 757 GUYGRILLIDPPT--TPVIEDIINTVIPEKFT 788

```
RESULT 6
A:Accession: A41341
C:Species: Bacillus subtilis
C:Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 09-Jul-2004
C:Accession: A41341; B41341; S39700; D69730
C:Stoma, A.; Ruto Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
J. Bacteriol. 173, 6889-6895, 1991
A:Title: Cloning and characterization of the gene for an additional extracellular serine
A:Reference number: A41341; MUID:92041574; PMID:1938892
A:Accession: A41341
A:Molecule type: DNA
A:Residues: 1-806 <SLD>
A:Cross-references: UNIPROT:P29141; UNIPARC:UPI000060C20; GB:M76590; NID:G143819; PIDN:
A:Accession: B41341
A:Molecule type: protein
A:Residues: 161-195 <SLD>
A:Cross-references: UNIPARC:UPI0000178D67
R:Glaeser, P.; Kunst, F.; Arnaut, M.; Coudart, M.E.; Gonzales, W.; Hullo, M.F.; Ionescu,
A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. 10, 371-384, 1993
A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
A:Reference number: S39655; MUID:95020537; PMID:7934828
A:Accession: S39700
A:Molecule type: DNA
A:Residues: 1-806 <GLA>
A:Cross-references: UNIPARC:UPI000060C20; EMBL:X73124; NID:G413923; PIDN:CAA51601.1; PI
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beter
C.; Bron, S.; Brouillet, S.; Bruchet, C.V.; Caldwell, B.; Capuani, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lamber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogilawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetalle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schaefer, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69730
A:Structure: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-806 <KUN>
A:Cross-references: UNIPARC:UPI000060C20; GB:Z99123; GB:AL009126; NID:G2636240; PIDN:CA
A:Experimental source: strain 168
C:Comment: The amino terminal sequence of the mature protein and a molecular weight of c
C:Genetics:
A:Gene: vpr
A:Start codon: TTG
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-160/Domain: propeptide #status predicted <PRO>
F:160-548/Domain: subtilisin homology #status atypical <SBT>
Query Match 9.3%; Score 286; DB 2; Length 806;
Best Local Similarity 22.5%; Pred. No. 3e-12;
Matches 132; Conservative 60; Mismatches 188; Indels 206; Gaps 19;
```

```
244 -----NDSSHEAFRGKITALLYALGRNNANDNGGHTYAGSVLNGXTNKGMAP 294
205 QYKGYDFVNDYDPRKETPTG-----DPRGEATDGTGTVAAANG-TIKGVAP 252
295 QANTVFGSINDSXXGGLGSPNLQTLFPOAXSAGARHTNWSGAAVNGAYTTDSRNVDDY 354
253 DATLATRVLPFG--SGTTENVLAGYRAVQDADVNNLSLNSLNPDMATSTAL-DW 309
355 VRKNDMTILFPAAGNEXPNGTISAPGTAKNAITVGATE---NLRFSGSY----- 401
310 AMSGVAVAVTNGSGNPGWGTVSGPSTRBSAIVGATQPLANEVAVTFGSSAKWNGYN 369
402 -----ADNI----- 405
370 KEDDVKALNKEVELVENGIGEAQDFEGKDLTGKVAVYKGSIAFVDKADNAKAGAGM 429
406 -----NHVAQF 411
430 VYVNNLSGEIRANVPCHSVPTIKSLDEKTVSALRAGETKTFKLTVSQALGEVADF 489
412 SSRGPTKD-GRIKEDVNAFGTXILSARSSLPADSSFWANHDSKYAYNGTSMATPIVAGN 470
490 SSRGPTVMDMMIKPDISAPGVNIVSTIPTHDPD-----HPYGVSKQGTSMASPHIAGA 543
471 VQALREHFVKKRGITTPPSL--LKAALIAQA---DXGLGPRNGNGMKRVTLDKSLNV 524
544 VAVIKQ-----AKPKVSVEQIKALIMNTAVTLKDSDBGVYFPHNAQAGSARI---MNA 593
525 AYVNESSXLSSTGKATYXFTTAGKPLKISLVMSDAPASTYASTL 570
594 --TKADSLVSGSYSTGYTFLEKNGENETKNETFTIENOSIRKSYTL 637
RESULT 7
525835
subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain TA41)
C:Species: Bacillus sp.
C:Date: 22-Nov-1993 #sequence revision 20-Feb-1995 #text_change 05-Oct-2004
C:Accession: S25835
R:Davall, S.; Feller, G.; Narinx, E.; Gerday, C.
Gene 119, 143-144, 1992
A:Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bacillu
A:Reference number: S25835; MUID:95012966; PMID:1358082
A:Accession: S25835
A:Molecule type: DNA
A:Residues: 1-419 <DAV>
A:Cross-references: UNIPARC:UPI0000BB77A; EMBL:X63533; NID:G40198; PIDN:CAA45096.1; PID
C:Superfamily: Subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-110/Domain: propeptide #status predicted <PRO>
F:111-419/Product: microbial serine proteinase #status predicted <MAT>
F:135-373/Domain: subtilisin homology <SBT>
F:144,184,359/Active site: Asp, His, Ser #status predicted
Query Match 9.1%; Score 279.5; DB 1; Length 419;
Best Local Similarity 29.6%; Pred. No. 3.5e-12;
Matches 97; Conservative 47; Mismatches 123; Indels 61; Gaps 16;
```

```
172 QXNKEVQLRGIEIXIAQXXSNDVYITAKPEYKMNVA-----RGIVKADVAQSS 222
81 QNNKNLVEKVPBL-----EITATNKPBL-ALYNMAASQSTPGWIKALYNNSTLST 132
223 YGLYGQGIIVAVADTGLDGRNDSNMH-EAFRGKITALLYALGRT---NNANDTNGHTV 278
133 SG--GAQINIAVADTGVNTNHPDLSNNVBOCKD-----FTYGNFTDNSCSDRGHGTHV 185
279 AGSVTLNGXTNKK--GNAPOANLVFOSIM--DSXGIGLGLPSNLQTLFPOAXSAGARIHT 333
186 AGSALANGGTGSGYGVAVPEADLWAVYVLGDDSGYVADIAEALRHAGDQATLNTKTVI 245
334 N-SWGAIVNGAYTTDSRNVDDYVAKNDMTILFPAAGNXPNGGTISAPGTAKNAITVGATE 392
```

Db 246 NMSLSSGSSSLIT---NAVDAYDKGVLLIAAGNSGKPPSGVFGALVNAVVALE 302
 Qy 339 NLRPFGSADNINVAQSSRGPTDG-----RIKPVMAFGTYILSARSLADDSF 446
 Db 303 NTON-GTY-----RAVDSSSRGKKTAGDYVIOKGDVEISAPGAIVST----- 346
 Qy 447 WANHDSKYVMGTSNATPIVAGNVQOL 474
 Db 347 W--PDGGYATISGTSMSAPHAAGLAANKI 372

RESULT 8

subtilisin-type protease (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
 C/Species: Bacillus halodurans
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C/Accession: G83753
 R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A/Reference number: AB3650; MUID:2051282; PMID:11058132
 A/Accession: G83753
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-799 <STO>
 A/Cross-references: UNIPROT:Q9KEM1; UNIPARC:UPI00000C39DC; GB:AP001510; GB:BA000004; NID
 A/Experimental source: strain C-125
 A/Genetics:
 A:Gene: vpr
 C/Superfamily: microbial serine proteinase vpr; subtilisin homology
 C/Keywords: hydrolase; serine proteinase
 F/1-29/Domain: signal sequence #status predicted <SIG>

Query Match 9.0%; Score 276.5; DB 2; Length 799;
 Best Local Similarity 23.5%; Pred. No. 1.4e-11;
 Matches 149; Conservative 61; Mismatches 198; Indels 227; Gaps 25;

Qy 78 LETVANNKLTAKXQNGP-ILR-----ETKQLEKTKAKILDYIPYATVEYE 125
 Db 50 LETV-----IVIEDPSIIIAKIQGQKSKHEKQARQSVIEQIDLVF--SSTVTHE 100
 Qy 126 GDVAXXXXXIIEHVSVEPTLKYXIDPOLFTGSAKXIVAKALDTRKQNKKEQLAGIEKI 185
 Db 101 YDFLFGSALR-----LPAHQIPSLIGDGHAVVFNIEEYETPDEY----- 144
 Qy 186 AQXXXXNDVXYITAKPEYK-VANQVARGIVKADVAQSSYGLYGQGIYAVADTGLDTSN 244
 Db 145 -----VIEKDAISPEMLDSAPFGANDAMEAGY--TGEETVAILDITGVDTYHP 191
 Qy 245 DSSMHEAFRGKITLALYALGRTNNANDT-----NGHGTIVAGSVLGNKXTNGMAPO 295
 Db 192 D--LVHAE-GDYKWDPIDNDDPQETPPGDRGIETTHGTHVAGVANAQLI-KGVA PD 247
 Qy 296 ANLVPGQIMDSXGIGLPSNLQTL--FSQAKSAGARHTMSKGAIVAGAYTTDSRNVDD 353
 Db 248 ANLLAYRL-----GGGSGSTAGVAGIERAQQDADIMNLSLGTLLNDPDPATIAL-D 302
 Qy 354 YVRKNDMTILPAGGEXPNGGTISAPGAKNAITVGAATENLRPSGSY-----ADNTNH 407
 Db 303 WAMAGVAVTNSNGSGRNNTVSGPISRDALISVGAAT--RLPYNKTKASVFTSDGIDY 359
 Qy 408 ----- 407
 Db 360 PSADIMGFPSDELELDEGETEYVAFAGLGRPDGFEVGVDEGKIALIVRGIFVEKXEN 419
 Qy 408 ----- 407
 Db 420 AKAGAVGAILYNNVAGVQPTVPGAIPITIMLSNEDGLKMRNELNGONTVTFSIIEPKL 479
 Qy 408 -----VAQSSRGPT-KDGRIKPDVMAFGTYILSARSLADDSFMANHDSKYVMGTSM 462
 Db 480 VGETVADSSSKRPVWHTMIKPDVSAFGVAVISTIPHQDDPY-----GYSGHQGTSM 533

Qy 463 APTIVAGNVQOLREHFVNKRGITPRKSLIKALINGAA---DXGLGYNGNGRCRVTL 518
 Db 534 ASPHYAGAAALILEKH-PWGV---DHVKALMTAENTLVDENGNRPHTQAG----- 584
 Qy 519 DKSINVAAYVNESSXLSQKATY-XFTATAGPLK 552
 Db 585 --SIRIVDAIBSETLVTPGSHSFGFTTKERKQVE 617

RESULT 9

subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
 C/Species: Bacillus sp.
 C/Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Oct-2004
 C/Accession: S23407
 R.Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
 Biochim. Biophys. Acta 1131, 111-113, 1992
 A>Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
 A/Reference number: S23407; MUID:92256481; PMID:1581352
 A/Accession: S23407
 A/Molecule type: DNA
 A/Residues: 1-420 <NAB>
 A/Cross-references: UNIPROT:P28842; UNIPARC:UPI0000136183; EMBL:X62369; NID:G40200; PIDN
 A/Genetics:
 A:Gene: sub1
 C/Superfamily: Subtilisin, subtilisin homology
 C/Keywords: extracellular protein; hydrolase; serine proteinase
 F/1-23/Domain: signal sequence #status predicted <SIG>
 F/24-111/Domain: propeptide #status predicted <PRO>
 F/112-420/Product: microbial serine proteinase #status predicted <MAT>
 F/116-374/Domain: subtilisin homology <SBT>
 F/145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 9.0%; Score 274.5; DB 1; Length 420;
 Best Local Similarity 28.3%; Pred. No. 7.9e-12;
 Matches 107; Conservative 56; Mismatches 142; Indels 73; Gaps 21;

Qy 168 LDTKQXN--KEVQLRGIEIXIAQXXXNDVXYITAKPE--YKVM--NDVARGIV---KA 216
 Db 73 MNEKQFNLKKNKNTLVKVPDL-----EIAVATDKPEALYNAMASQSTPMGIKATYNN 128
 Qy 217 DVAQSYGLYQGGQVAVADTGLDTRNDSSMHEAFRGKITRL-YALGRT---NNADT 271
 Db 129 SITQTS---GGGGINIAVLDTGVNTN-----HPDLRNNVQCKDFTVGTYYTNNSCTDR 179
 Qy 272 NGHGTIVAGSVLGNKXTNK-----GMAPOANLVQSIM--DSXGIGLGLPSNLQTLFSQXS 326
 Db 180 QGHGTHVAGSALADGCTGNGYGVAPDADLMAKYKVLGDGSGSYADDDIAALIHAGDQATA 239
 Qy 327 AGARITHTN-SWGAAVNGAYTTDSRNVDDYVRKNDTILFAAGNEXPNGGTISAPGTAKNA 385
 Db 240 LNTKVLVIMSIGSSSESLINAVN--YSNKGVLIIAAGNSGPGYSGIYFGALVNA 296
 Qy 386 ITVGATENTLRPSFGSYADNINVAQSSRGPT-KDG-----RIKPDVMAFGTYILSARSS 439
 Db 297 VAVVALLEN-KYENGTY-----RVADFSRGYSWTDGVAIQKGDVEISAPGAAYIST--- 347
 Qy 440 LAPDSSFPANHDSKYVMGTSNATPIVAGNVQOLREHFVNKRGITPRKSL-----L 491
 Db 348 -----W--PDGGYATISGTSMSAPHAAGLAANKIWAQPSASNDVAGELQYRAYENDI 398
 Qy 492 KAALTAGAAD---XGLGY 506
 Db 399 LSGYVAGYGDGDPFASGFGF 416

RESULT 10

S11890
 serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca.
 N/Alternate names: subtilisin-related proteinase
 C/Species: Xanthomonas campestris pv. campestris
 C/Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 31-Dec-2004

C|Accession: S11890
R|Lin, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A|Title: A multipurpose broad host range cloning vector and its use to characterise an
A|Reference number: S11890; MUID:90251253; PMID:2187155
A|Accession: S11890
A|Molecule type: DNA
A|Residues: 1-580 <LTI>
A|Cross-references: UNIPROT:P23314; UNIPARC:UP1000012A398; EMBL:X51635; NID:g48533; PIDN
A|Experimental source: Xanthomonas campestris pv. campestris
A|Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-ALA
C|Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
C|Keywords: extracellular protein; hydrolase; serine protease
F|1-32/Domain: signal sequence #status predicted <SIG>
F|168-423/Domain: subtilisin homology <SBT>

Query Match 8.8%; Score 270; DB 2; Length 580;
Best Local Similarity 25.0%; Pred. No. 2.5e-11;
Matches 135; Conservative 66; Mismatches 194; Indels 146; Gaps 26;

QY 162 LVKA-XALDTKQXKEVQLRGEXIAQXXXSNDVXYITAKP-EYKVMNDVARGIVKA--D 217
DB 99 LVRAALDRABAEFLMRQLAADPNVQSVENVQILHATLTLPDTRLSEQMAFGTTNAGLN 158
QY 218 VAQSSYGLYGGQGVAVVADTG----DTGRDSSMHEAFRGKITLALYALGRTNAND--- 270
DB 159 IRPAMDKATGSGTVAVVADTGITSHADLNANLADYDPTSDATTVARDGNGSDSNADGCD 218
QY 271 -----TNGHTVAGSV--LGNGKTN-KGMAPOANLVEFQSIMDSXG--- 309
DB 219 WYANECGAGTPAASSMHGTHVAGTVAATNNTTGVAGTAVGAKVVEVRLGKCGSL 278
QY 310 -----LGLPSTNQ--TLFQOASAGARHITNSKGAALVNGAYTTDSRNVDY 354
DB 279 DIADALVWASGQTVSGIPANANPAEVINMSLGCGSGSCSTWQON-AINGAVSRGT----- 331
QY 355 VRKNDMTLFAAGNEXPNRGITISAPGTAKNAITVATEN--LRPSFGSYADININVAQFS 412
DB 332 -----TVVAAGNDASVNSG-SLPANCANVAVAAATTSAGAKASVSNFGTGI----- 377
QY 413 SRGPKDRIKPDVWAPGTXTLSARSS--LAPDSSFMANHDSKYAVMGSTSMATPIVAGN 470
DB 378 -----DVSAPSSSILSTINSGTTTSGS-----ASYASVNGTSMASPHVAGV 418
QY 471 VAQREHFVKNRGITPK--PSLTK--AALLAGADXLGY-----PENGQ 511
DB 419 VALVQS--VAPPTALPPAVERETLKNTARALPGACGCGAGIVNADAAVTAINGSGSG 476
QY 512 GMRGVTLDKSLNVAVYVNSSKXSTGQKATYKFTATAGKPKISLVWSDAPASTTASVTL- 570
DB 477 GGGGNTLTNGTPTVGLG---AATGAELNYTITVPAG-----SGTLVTTTS 518
QY 571 --VNDLDELVI---TAPNGTYYVGNDFXXPKXXNMMDGRNVENVFINKPQSGTYTIEVOAY 625
DB 519 GSGGADALYVAGSAPTSATYCRPYRS-----GNAETCTITAP-SGTYYVRLKAY 568
QY 626 N 626
DB 569 S 569

RESULT 11
JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N|Alternate names: subtilase
C|Species: Alteromonas sp.
C|Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 05-Oct-2004
C|Accession: JC4908
R|Tanjilo, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Inada, C.; Okami, Y.; Inamori, Y.
Biochem. Biotechnol. Biochem. 60, 1284-1288, 1996
A|Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacter
A|Reference number: JC4908; MUID:97141200; PMID:9897544
A|Accession: JC4908

A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-715 <TSU>
A|Cross-references: UNIPROT:P70765; UNIPARC:UP1000017A928; DDBJ:D38600; NID:g1536787; PI
A|Experimental source: strain O-7
A|Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sens
C|Genetics:
A|Gene: aprI
C|Superfamily: extracellular alkaline serine protease precursor, Alteromonas type
C|Keywords: hydrolase
F|1-40/Domain: signal sequence #status predicted <SIG>
F|41-150/Domain: amino-terminal propeptide #status predicted <AP>
F|151-486/Product: alkaline serine protease I #status predicted <MAT>
F|182-452/Domain: subtilisin homology <SBT>
F|497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F|239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 8.8%; Score 268.5; DB 2; Length 715;
Best Local Similarity 25.8%; Pred. No. 4.3e-11;
Matches 142; Conservative 51; Mismatches 201; Indels 157; Gaps 28;

QY 169 DTKQXKEVQLRGEXIAQXXXSNDVXYIT-----AKREYKMD-----VARGIV 214
DB 117 DKLSAKBAQAFMEVVA-----SGNVEYIEIDQMLKPFATPNDRYNDQWHTYEAAAGI- 171
QY 215 KADVAQSSYGLYGGQGVAVVADTG----LDTGRDSSMHEAFRGKITLALYALGRTNAND 270
DB 172 NAPAAMDK--ATGGGVVAVVADTGYPRLDLDANILPGYDMSITFVANDGARDNDARD 229
QY 271 -----TNG-----HGTAVAG--SVLGNGXTNKGMAPOANLVEFQSI 302
DB 230 PGDAVTRGECGTDSGQVPVPAADODSSMHGTHVAGTVAATYNNGBGVAVYDAKVPVR 289
QY 303 IMDSXGGLGGLPSNLQTLFQOASAGARHITNSKGAALV-----NGAYTTDSRNVDY 355
DB 290 VL---GKCGGLTSDIADGIWASGSDRVPANANPAVIVNMSLGCGACSAKTQNALINQA 346
QY 356 RKNDMTLFAAGNEXPNRGITISAPGTAKNAITVAT--ENLRPSFGSYADININVAQFS 413
DB 347 RNNGTIVIVAGNNDNSANYN--PGNCGVNVVASVGRDSRAYSYNGAI----- 397
QY 414 RGPDKRIKPDVWAPGTXTLSARSSLAPDSFPANHDS-----KYAVMGSTSMATP 465
DB 398 -----DVAAPG---GAQSPADDEPGILSTNMSGAGPANSYHSQSTSWAAP 442
QY 466 IVAGNVAGLR-----EHFVKN--RGITPKSLKRALI--AGADXLGY---Y 506
DB 443 HVAGVAAALIKQAKXSATPDEVETILKNTTSFAGSCNCGGVDDAAAVVEALGDVVT 502
QY 507 PNGQGMGRVTLDKSLNVAVYVNSSKXSTGQKATYKFTATAGKPKISLVWSDAPASTTA 566
DB 503 PTGN-----TLED--GVAKTGLSGAAGSNQ--PFTDVPKGR-----TNV 538
QY 567 SVTL---VNDLDELVI---TAPNGTYYVGNDFXXPKXXNMMDGRNVENVFINKPQSGTYT 619
DB 539 TPTWSGGTGADLDLYK-----LGSQ---PTSSSYOCRPEEGNALVCSTDAQACGYH 588
QY 620 IEQAVNVVPG 630
DB 589 VMINGYKAYSG 599

RESULT 12
F69730
cell wall-associated protein precursor wprA - Bacillus subtilis
C|Species: Bacillus subtilis
C|Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C|Accession: F69730
R|Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerzon, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A|Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler

F:118-117/Domains: propetidein #status predicted <PRO>
F:123-408/Product: aquaglysin I #status experimental <MAT>
F:117-164/Domains: subtilisin homology <SBT>
F:425-257,281-283/Region: SI specificity crevice #status predicted
F:105-513/Domains: carboxyl-terminal propetidein #status predicted <CPR>
F:166,197,349/Active site: Asp, His, Ser #status predicted

[illegible]

Search completed: April 7, 2006, 17:43:35
Job time : 29.5215 Secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 17:33:06 ; Search time 149.117 Seconds

(without alignments)
3028.090 Million cell updates/sec

Title: US-10-784-870-2

Perfect score: 3059
Sequence: 1 MEXKKVFLSVLSAAALIST.....EVQAVNPVGPQXFSLATVN 640

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3029	99.0	640	Q93UV9_9BACI	Q93UV9 bacillus sp
2	3020	98.7	639	Q9AQR3_9BACI	Q9AQR3 bacillus sp
3	3010	98.4	640	Q76L84_9BACI	Q76L84 bacillus sp
4	2082	68.1	433	Q9AQR0_9BACI	Q9AQR0 bacillus sp
5	1952.5	63.8	433	Q9AQR1_9BACI	Q9AQR1 bacillus sp
6	1948.5	63.7	433	Q9AQR4_9BACI	Q9AQR4 bacillus sp
7	1941.5	63.5	433	Q9AQR2_9BACI	Q9AQR2 bacillus sp
8	863.5	28.2	697	Q4NB18_9M1CC	Q4NB18 arthrobacte
9	793	25.9	1748	Q4NWB5_9DBBL	Q4NWB5 anaeromyxob
10	786.5	25.7	711	Q6MKR4_BDBBA	Q6MKR4 bdellovibri
11	638	20.9	2030	Q747P6_GBOSL	Q747P6 geobacter s
12	594	19.4	1088	Q4HUY5_GIBZE	Q4HUY5 gibberella
13	526.5	17.2	1741	Q54M84_DICDI	Q54M84 dictyostell
14	518.5	16.9	1743	TACQ_DICDI	Q23668 dictyostell
15	493	16.1	1825	Q8T9W1_DICDI	Q8T9W1 dictyostell
16	478	15.6	1752	Q9GTN7_DICDI	Q9GTN7 dictyostell
17	477.5	15.6	1905	TAGB_DICDI	P54683 dictyostell
18	477.5	15.6	1906	Q54M83_DICDI	Q54M83 dictyostell
19	421.5	13.8	1388	Q580L9_9TRYX	Q580L9 trypanosoma
20	413	13.5	663	Q5U1Z5_PYKCO	Q5U1Z5 pyrococcus
21	412	13.3	561	Q8RBJ2_THRTN	Q8RBJ2 thermotanaer
22	408	13.3	654	Q8UBC9_PYRFU	Q8UBC9 pyrococcus
23	391	12.8	1239	Q9FBZ4_STRCO	Q9FBZ4 streptomyce
24	375	12.3	430	Q8ENV1_OCEIH	Q8ENV1 oceanobacill
25	368.5	12.0	1253	Q9FCO6_STRCO	Q9FCO6 streptomyce
26	356	11.6	1102	P95684_STRAW	P95684 streptomyce
27	352.5	11.5	1208	Q82B14_STRAW	Q82B14 streptomyce
28	343.5	11.2	444	Q9KBJ7_BACDH	Q9KBJ7 bacillus ha
29	340	11.1	1105	Q82I39_STRAW	Q82I39 streptomyce
30	337	11.0	1245	Q8KHB6_STRVD	Q8KHB6 streptomyce
31				Q9RL54_STRCO	Q9RL54 streptomyce

ALIGNMENTS

32	336.5	11.0	412	2	Q9AER6_THERMO	Q9AER6 thermotanaer
33	336.5	11.0	412	2	Q8RC68_THRTN	Q8RC68 thermotanaer
34	335	11.0	795	2	Q5NM24_9ARCH	Q5NM24 uncultured
35	333	10.9	1237	2	Q8G9T4_STRAZ	Q8G9T4 streptomyce
36	332	10.9	442	2	Q5L315_GBOKA	Q5L315 geobacillus
37	332	10.9	442	2	Q5L315_GBOKA	Q5L315 geobacillus
38	332	10.6	1294	2	Q5OHM7_STRSH	Q5OHM7 streptomyce
39	323	10.6	1899	2	Q67RJO_STWTH	Q67RJO symbiodacte
40	321.5	10.5	1220	2	Q9LOA0_STRCO	Q9LOA0 streptomyce
41	319.5	10.4	435	2	Q8EMJ3_OCEIH	Q8EMJ3 oceanobacill
42	319.5	10.4	442	2	Q31788_BACSU	Q31788 bacillus su
43	318.5	10.4	524	2	Q5JEB9_PYKCO	Q5JEB9 pyrococcus
44	315	10.3	818	2	Q79CG2_BACSP	Q79CG2 bacillus sp
45	315	10.3	824	2	Q45464_BACSP	Q45464 bacillus sp

RESULT 1	Q93UV9_9BACI	PRELIMINARY;	PRT;	640 AA.
ID	Q93UV9_9BACI			
AC	Q93UV9_9BACI			
DT	01-DEC-2001 (TREMblrel. 19, Created)			
DT	01-OCT-2002 (TREMblrel. 22, Last sequence update)			
DT	01-MAR-2004 (TREMblrel. 26, Last annotation update)			
DE	Name=PROF;			
GN	Protease.			
OS	Bacillus sp. KSM-KP43.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_Taxid=109322;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=KP43;			
RA	Itoh S., Saeki K.;			
RL	Submitted (NOV-2000) to the EMBL/Genbank/DDBJ databases.			
DR	EMBL; AB051423; BAB55674.2; -; Genomic DNA.			
DR	PDB; 1WMD; X-ray; A=207-640.			
DR	PDB; 1WME; X-ray; A=207-640.			
DR	PDB; 1WME; X-ray; A=207-640.			
DR	MEROPS; S08.123; -;			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0004288; F:subtilase activity; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR007280; Pept_Bact_C.			
DR	InterPro; IPR002099; Pept_S8_S53.			
DR	Pfam; PF00082; Peptidase_S8_1.			
DR	Pfam; PF04151; PPC; 1.			
DR	PRINTS; PR00723; SUBTILISIN.			
DR	PROSITE; PS00137; SUBTILASR_HIS; 1.			
DR	PROSITE; PS00138; SUBTILASR_SER; 1.			
KW	Hydrolase; Serine protease.			
SQ	SEQUENCE 640 AA; 67991 MW; 48BAF77BE9D592C15 CRC64;			

Query Match	99.0%; Score 3029; DB 2; Length 640;
Best Local Similarity	93.3%; Pred. No. 3.2e-210;
Matches 597; Conservative	0; Mismatches 43; Indels 0; Gaps 0;

QY	1	MEXKKVFLSVLSAAALISTVALNPSAGAXPFLDRKGIOTTTDDXGFSKXOTGA	60
DB	1	MEXKKVFLSVLSAAALISTVALNPSAGAXPFLDRKGIOTTTDDXGFSKXOTGA	60
QY	61	FLVSEENVKLKGLKKLETPANNKKAIXFNGPILBETKQXLBXTGAKTIDYIPDY	120
DB	61	FLVSEENVKLKGLKKLETPANNKKAIXFNGPILBETKQXLBXTGAKTIDYIPDY	120
QY	121	IYETGIDVYKSKKXIEHVSVEPIYKXIDPQLFTKASALVAKKALDTQXNREV	180
DB	121	IYETGIDVYKSKKXIEHVSVEPIYKXIDPQLFTKASALVAKKALDTQXNREV	180
QY	181	GIEKXIAKKKXNDVYITRAKPEYKMNVDAGIVADVAQSSYGLYGGQIVAVD	240
DB	181	GIEKXIAKKKXNDVYITRAKPEYKMNVDAGIVADVAQSSYGLYGGQIVAVD	240

```

QY 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVE 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVE 300
QY 301 QSIMDSXGGLGGLPSNLQTLFSAQXSAGARLHTNSGAAVNGAYTTDSRNVDVYRKNDM 360
DB 301 QSIMDSXGGLGGLPSNLQTLFSAQXSAGARLHTNSGAAVNGAYTTDSRNVDVYRKNDM 360
QY 361 TILFAAGNEXPNVGGITTSAPGTAKNAITVCATENLRPSFGSYADNINHYAQSSRGPTKDG 420
DB 361 TILFAAGNEXPNVGGITTSAPGTAKNAITVCATENLRPSFGSYADNINHYAQSSRGPTKDG 420
QY 421 RIKPVPVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFPK 480
DB 421 RIKPVPVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFPK 480
QY 481 NRGITPKPSLLKALJAGAADXGLGYPNGNGMGWRTLLDKSLNVAAYNNESSXLTSTOKAT 540
DB 481 NRGITPKPSLLKALJAGAADXGLGYPNGNGMGWRTLLDKSLNVAAYNNESSXLTSTOKAT 540
QY 541 YXFITAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPKXPKXKXND 600
DB 541 YXFITAGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPKXPKXKXND 600
QY 601 GRNNVENVPINPKPQSGTYTIEVOAYNVVPVGPQFSLAIVN 640
DB 601 GRNNVENVPINPKPQSGTYTIEVOAYNVVPVGPQFSLAIVN 640

```

RESULT 2

Q9AOR3_9BACI PRELIMINARY; PRT; 639 AA.

```

ID 09AOR3_9BACI PRELIMINARY; PRT; 639 AA.
AC 09AOR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protease.
GN Name:PROA;
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133778;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=9860;
RX MEDLINE=20568675; PubMed=11118284; DOI=10.1006/birc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046403; BAB21266.2; -; Genomic_DNA.
DR HSSP; P00782; 1SUP.
DR SMR; Q9AOR3; 206-639.
DR MEROPS; S08.123; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASB_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 639 AA; 68186 MW; 316AF6FFDB4FF54 CRC64;

```

Query Match 98.7%; Score 3020; DB 2; Length 639;
 Best Local Similarity 93.4%; Pred. No. 1.4e-209;
 Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

QY 4 KKKVFLSVLSAAAILSTVALXNPSAGXARFDDLPKGIQTITDDXGFSKQKQTGAALFLV 63
DB 3 KKKVFLSVLSAAAILSTVALXNPSAGXARFDDLPKGIQTITDDXGFSKQKQTGAALFLV 62
QY 64 ESENVKLLKGLXKKLETVPANMKLHIQFPNPILEETKQILEXTGAKILDYIPVAYIVE 123
DB 63 ESENVKLLKGLXKKLETVPANMKLHIQFPNPILEETKQILEXTGAKILDYIPVAYIVE 122
QY 124 YEGDVXSKXXXIENHVESIEPTLPKXIIDPOLFTGASLTVKXALDTPQXKVEVLRGIE 183
DB 123 YEGDVQSKVRSIEHVESIEPTLPKXIIDPOLFTGASLTVKXALDTPQXKVEVLRGIE 182
QY 184 XIAQXXSNDVXYITAKPEYKAMDVARGIYKADVAQSSYGLYGGQGVAAVADTGLDGR 243
DB 183 EIAQTVASNDVHYITAKPEYKAMDVARGIYKADVAQSSYGLYGGQGVAAVADTGLDGR 242
QY 244 NDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVEQSI 303
DB 243 NDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLNGKXTNKGMAPOANLVEQSI 302
QY 304 MDSXGGLGGLPSNLQTLFSAQXSAGARLHTNSGAAVNGAYTTDSRNVDVYRKNDMTIL 363
DB 303 MDSXGGLGGLPSNLQTLFSAQXSAGARLHTNSGAAVNGAYTTDSRNVDVYRKNDMTIL 362
QY 364 FAAGNEXPNVGGITTSAPGTAKNAITVCATENLRPSFGSYADNINHYAQSSRGPTKDGRIK 423
DB 363 FAAGNERPNVGGITTSAPGTAKNAITVCATENLRPSFGSYADNINHYAQSSRGPTKDGRIK 422
QY 424 PDVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFPKXRG 483
DB 423 PDVMAPGTITLSARSSLPDSSFWANHDSKYAYMGSTSMATPIVAGNVAQLREHFPKXRG 482
QY 484 ITPRPSLLKALJAGAADXGLGYPNGNGMGWRTLLDKSLNVAAYNNESSXLTSTOKATYXP 543
DB 483 ITPRPSLLKALJAGAADXGLGYPNGNGMGWRTLLDKSLNVAAYNNESSXLTSTOKATYXP 542
QY 544 TATGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPKXPKXKXNDGRN 603
DB 543 TATGKPLKISLVMSDAPASTTASVTLVNDLVLITAPNGTYVGNDFPKXPKXKXNDGRN 602
QY 604 NVENVPINPKPQSGTYTIEVOAYNVVPVGPQFSLAIVN 640
DB 603 NVENVPINPKPQSGTYTIEVOAYNVVPVGPQFSLAIVN 639

```

RESULT 3

Q76L84_9BACI PRELIMINARY; PRT; 640 AA.

```

ID Q76L84_9BACI PRELIMINARY; PRT; 640 AA.
AC Q76L84;
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Protease.
GN Bacillus sp. KSM-9865.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=192495;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-9865;
RA Okuda M., Saeki K., Kobayashi T.;
RT Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
RL EMBL; AB084155; BAC82522.1; -; Genomic_DNA.
DR HSSP; P00782; 1AON.
DR SMR; Q76L84; 207-640.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILASB_HIS; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.

```

DR PROSITE; PS00138; SUBTILASE_SBR; 1.
 KM Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 640 AA; 67940 MW; 42178150FP28688F CRC64;
 Query Match 98.4%; Score 3010; DB 2; Length 640;
 Best Local Similarity 92.8%; Pred. No. 7.6e-209;
 Matches 594; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 MEXKKKVPISVSAALISTVALXNPAGAKXFDLDFGIQTITDXXKQXQTGAAA 60
 DB 1 MEKKKKVPLSVSAALISTVALSNPSAGARFDFDIFGIQTITDXXKQXQTGAAA 60
 QY 61 FLVESBNVTLXKGLKQLETPVANNGLHXQFNGPILSEKTKXLBXTGAKITDIYIPDVA 120
 DB 61 FLVESBNVTLXKGLKQLETPVANNGLHXQFNGPILSEKTKXLBXTGAKITDIYIPDVA 120
 QY 121 IVEYEGDVAXXXXXXIEHVESVEPYLPKXXIDPOLFTKGASXLVKAXALDTKXNKVQLR 180
 DB 121 IVEYEGDVAXSATSTIEDVSVEPYLPYRIDPOLFTKGASXLVKAXALDTKXNKVQLR 180
 QY 181 GIEKTLAOKXXXSDVXYITNAKPEYKVMNDVARGIVKADVAQSSYGLYGQGIYVADTGLD 240
 DB 181 GIEQIQAQFATSDVYITNAKPEKVMNDVARGIVKADVAQSSYGLYGQGIYVADTGLD 240
 QY 241 TGRNDSMHEAFRGKITLALGRTNANDTNGHGHVAGSVLGNQXTKGMAPQANLYF 300
 DB 241 TGRNDSMHEAFRGKITLALGRTNANDTNGHGHVAGSVLGNQXTKGMAPQANLYF 300
 QY 301 QSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDM 360
 DB 301 QSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDM 360
 QY 361 TILPAAAGNEKPGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
 DB 361 TILPAAAGNEKPGGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
 QY 421 RIKPVPVAPGTIILSARSLAPDSSFWAHDSKLYVMGTSMTATPIVAGNVQLREHPFK 480
 DB 421 RIKPVPVAPGTIILSARSLAPDSSFWAHDSKLYVMGTSMTATPIVAGNVQLREHPFK 480
 QY 481 NRGITPKPSLKAALIAAGADXLGYPNGNGMGRTYLDKSLNVAAYNNESSXLSSTOKAT 540
 DB 481 NRGITPKPSLKAALIAAGADXLGYPNGNGMGRTYLDKSLNVAAYNNESSXLSSTOKAT 540
 QY 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLIVITAPNGTYVGNDFPKXPKXKND 600
 DB 541 YSFTATAGKPLKISLWSDAPASTTASVTLVNDLIVITAPNGTYVGNDFPKXPKXKND 600
 QY 601 GNNNVENPINKPQSTYTIIEVOAYNVPGPQFSLAIYN 640
 DB 601 GNNNVENPINKPQSTYTIIEVOAYNVPGPQFSLAIYN 640

RESULT 4
 Q9AQR0_9BACT PRELIMINARY; PRT; 434 AA.
 AC Q9AQR0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Protease (Fragment).
 GN Name=PROB;
 OS Bacillus sp. NV1.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=133781;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=NV1;
 RX MEDLINE=2058675; PubMed=1118284; DOI=10.1006/birc.2000.3931;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 alkaliophilic Bacillus spp.: enzymatic properties, sequences, and

RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 DR EMBL; AB046406; BAB21265.1; -, Genomic_DNA.
 DR HSSP; P00782; ISUP.
 DR SMR; Q9AQR0; 1-434.
 DR MEROPS; S08.123; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR007280; Pept_Bact_C.
 DR InterPro; IPR000209; Pept_S8_S53.
 DR Pfam; PF04151; Ppc; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Hydrolyase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;
 Query Match 68.1%; Score 2082; DB 2; Length 434;
 Best Local Similarity 91.5%; Pred. No. 5.2e-142;
 Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 207 NDVARGIKADVAQSSYGLYGQGIYVADTGLDTRGNDSSMHEAFRGKITLALGRTN 266
 DB 1 NDVARGIKADVAQSSYGLYGQGIYVADTGLDTRGNDSSMHEAFRGKITLALGRTN 266
 QY 267 NANDNGHGHVAGSVLGNQXTKGMAPQANLYFQSIMDSXGGLGGLPSNLQTLFQOAXS 326
 DB 61 NANDNGHGHVAGSVLGNQXTKGMAPQANLYFQSIMDSXGGLGGLPSNLQTLFQOAXS 326
 QY 327 AGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILPAAAGNEKPGGTISAPGTAKNAI 386
 DB 121 AGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILPAAAGNEKPGGTISAPGTAKNAI 386
 QY 387 TYGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKQDVMAPGTIIISARSLAPDSSF 446
 DB 181 TYGATENLRPSFGSYADNINHVAQFSSRGPTDGRIKQDVMAPGTIIISARSLAPDSSF 446
 QY 447 WANHDSKTYAVMGTSMTATPIVAGNVQOLREHPVKRGTTPKPSLKAALIAAGADXLGY 506
 DB 241 WANHDSKTYAVMGTSMTATPIVAGNVQOLREHPVKRGTTPKPSLKAALIAAGADXLGY 506
 QY 507 PNGNGMGRTYLDKSLNVAAYNNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTT 566
 DB 301 PNGNGMGRTYLDKSLNVAAYNNESSXLSSTOKATYXFTATAGKPLKISLWSDAPASTT 566
 QY 361 SVTLVNDLIVITAPNGTYVGNDFPKXPKXKNDGRNVENPINKPQSTYTIIEVOAYN 626
 DB 361 SVTLVNDLIVITAPNGTYVGNDFPKXPKXKNDGRNVENPINKPQSTYTIIEVOAYN 626
 QY 627 VPVGPOFSLAIYN 640
 DB 421 VPVGPOFSLAIYN 640

RESULT 5
 Q9AQR1_9BACT PRELIMINARY; PRT; 433 AA.
 AC Q9AQR1;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Protease (Fragment).
 GN Name=PROD;
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=133780;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SD521;

RX MEDLINE=2056675; PubMed=1118284; DOI=10.1006/brc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RT Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic *Bacillus* sp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046405; BAB21268.1; -; Genomic_DNA.
DR HSSP; Q45670; IDBI.
DR SMR; Q9AQR4; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Pept_Bact_C.
DR InterPro; IPR000209; Pept_Bact_C.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18FE60DDC CRC64;

Query Match 63.8%; Score 1952.5; DB 2; Length 433;
Best Local Similarity 86.6%; Pred. No. 12e-132;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 207 NDVARGIKADVAAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 266
DB 1 NDVARGIKADVAAQNNYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 60
QY 267 NANTNGHGTIVAGSVLNGKTNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSQAXS 326
DB 61 NANTNGHGTIVAGSVLGN -ALNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSQAMN 119
QY 327 AGARHTNSMGAANVGAATYDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
DB 120 AGARHTNSMGAANVGAATYANSRQVDEYVRNNDMTILFAAGNEXPNGGTISAPGTAKNAI 179
QY 387 TVGATENLRPSFGSYADININVAQFSRSGPTDGRIRKEDVMAFGTXILSARSLAPDSF 446
DB 180 TVGATENLRPSFGSLADNPNIHQFSRSGATRDGIRKEDVTAPEGFTLSARSLAPDSF 239
QY 447 WANHSKYAAMGCTSMATPIVAGNVAQLREHFVKNRGITTPKPSLKAALAGATVGLGY 506
DB 240 WANYSKYAAMGCTSMATPIVAGNVAQLREHFVKNRGITTPKPSLKAALAGATVGLGY 299
QY 507 PNGNGMGRVTLDSGLNVAAYNESSXLSTQKATYXFTATAGKPLKISLVMSDAPASTTA 566
DB 300 PNGNGMGRVTLDSGLNVAAYNEATALTQKATYXFTQAGKPLKISLVMTDAPGSTTA 359
QY 567 SVTLVNDLDLVITAPNGTXYYGNDFFXPKXKNMGRNNVENVFINKPQSGTYTIEVQAYN 626
DB 360 SVTLVNDLDLVITAPNGKYVGNDFSYPDNMMGRNNVENVFINKPQSGTYTIEVQAYN 419
QY 627 VPVGPOKPSLAIVN 640
DB 420 VPSPQRPFLAIVH 433

RESULT 6
Q9AQR4_9BACT
ID Q9AQR4_9BACT PRELIMINARY; PRT; 433 AA.
AC Q9AQR4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Protease (Fragment).
GN Name=PROA;
OS *Bacillus* sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.

OX NCBI_TaxID=127889;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=D6;
RX MEDLINE=2056675; PubMed=1118284; DOI=10.1006/brc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RT Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic *Bacillus* sp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046402; BAB21265.1; -; Genomic_DNA.
DR HSSP; Q45670; IDBI.
DR SMR; Q9AQR4; 1-433.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR000209; Pept_Bact_C.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

Query Match 63.7%; Score 1948.5; DB 2; Length 433;
Best Local Similarity 86.4%; Pred. No. 2.3e-132;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIKADVAAQSSYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 266
DB 1 NDVARGIKADVAAQNNYGLYGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 60
QY 267 NANTNGHGTIVAGSVLNGKTNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSQAXS 326
DB 61 NANTNGHGTIVAGSVLGN -ALNKGMAPOANLVFQSIIMDSXGGLGSLPSNLQTLFSQAMN 119
QY 327 AGARHTNSMGAANVGAATYDTSRNVDDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
DB 120 AGARHTNSMGAANVGAATYANSRQVDEYVRNNDMTILFAAGNEXPNGGTISAPGTAKNAI 179
QY 387 TVGATENLRPSFGSYADININVAQFSRSGPTDGRIRKEDVMAFGTXILSARSLAPDSF 446
DB 180 TVGATENLRPSFGSLADNPNIHQFSRSGATRDGIRKEDVTAPEGFTLSARSLAPDSF 239
QY 447 WANHSKYAAMGCTSMATPIVAGNVAQLREHFVKNRGITTPKPSLKAALAGATVGLGY 506
DB 240 WANYSKYAAMGCTSMATPIVAGNVAQLREHFVKNRGITTPKPSLKAALAGATVGLGY 299
QY 507 PNGNGMGRVTLDSGLNVAAYNESSXLSTQKATYXFTATAGKPLKISLVMSDAPASTTA 566
DB 300 PNGNGMGRVTLDSGLNVAAYNEATALTQKATYXFTQAGKPLKISLVMTDAPGSTTA 359
QY 567 SVTLVNDLDLVITAPNGTXYYGNDFFXPKXKNMGRNNVENVFINKPQSGTYTIEVQAYN 626
DB 360 SVTLVNDLDLVITAPNGKYVGNDFSYPDNMMGRNNVENVFINKPQSGTYTIEVQAYN 419
QY 627 VPVGPOKPSLAIVN 640
DB 420 VPSPQRPFLAIVH 433

RESULT 7
Q9AQR2_9BACT
ID Q9AQR2_9BACT PRELIMINARY; PRT; 433 AA.
AC Q9AQR2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Protease (Fragment).
GN Name=PROC;
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_Taxid=133779;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RX MEDLINE=20568675; PubMed=1118284; DOI=10.1006/dbrc.2000.3931;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319 (2000).
DR EMBL; AB046404; BAB21267.1; -; Genomic_DNA.
DR HSSP; Q45670; 1DB1.
DR SMR; Q9AQR2; 1-433.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0004289; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR007280; Pept_Bact_C.
DR InterPro; IPR00209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; Ppc; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASIS_HIS; 1.
DR PROSITE; PS00138; SUBTILASIS_SER; 1.
KM Hydrolyase; Protease; Serine protease.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 433 AA, 45588 MW; B81291A803C775AE CRC64;

Query Match 63.5%; Score 1941.5; DB 2; Length 433;
Best Local Similarity 86.2%; Pred. No. 7.3e-132;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 207 NDVARGIKADVAOSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALYALGRTN 266
DB 1 NDVARGIKADVAQNNYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALYALGRTN 60
QY 267 NANTNGHGTAVAGSVLNGXGKTNKGAPOANLVPOSIMDSXGGLGGLPSNLQTLPSQAXS 326
DB 61 NASDPNGHGTAVAGSVLGN-ALNKGMAPOANLVPOSIMDSXGGLGGLPSNLQTLPSQAXN 119
QY 327 AGARIHTNSWGAANVAGATYDTSRNVDDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
DB 120 AGARIHTNSWGAANVAGATYDTSRNVDDVYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 179
QY 387 TVGATENTLRPFSGFADYNINHVAFSSRGFTYKGRIKPDVMAPGTYILSARSSILAPDSGF 446
DB 180 TVGATENTLRPFSGFADYNINHVAFSSRGFTYKGRIKPDVMAPGTYILSARSSILAPDSGF 239
QY 447 WANHSKYAVNGGSMATPIVAGNVAOLREHFVNRGITPKPSILKALIIAAGAXGIGY 506
DB 240 WANHSKYAVNGGSMATPIVAGNVAOLREHFVNRGITPKPSILKALIIAAGATVGLGY 299
QY 507 PNGNGWGRVTLNDSKLVAVYVNESKLSSTQKATYFTATAGPKLISLVMSDAPASTTA 566
DB 300 PNGNGWGRVTLNDSKLVAVYVNESKLSSTQKATYFTATAGPKLISLVMSDAPASTTA 359
QY 567 SVTLVWDDLVYITANGTYKYGNDYKPKXKXNDGRNVENYFIKXPOSGITTIIVQAVN 626
DB 360 SVTLVWDDLVYITANGTYKYGNDYKPKXKXNDGRNVENYFIKXPOSGITTIIVQAVN 419
QY 627 VPVGPQXSLAIYN 640
DB 420 VPVGPQXSLAIYN 433

RESULT 8
Q4NB18_9M1CC
ID Q4NB18_9M1CC PRELIMINARY; PRT: 697 AA.

AC Q4NB18;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, seedolisin.
GN ORFNames=ArthDRAP7_0589;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteri; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_Taxid=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAHG0100025; EAL94539.1; -; Genomic_DNA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASIS_HIS; 1.
DR PROSITE; PS00138; SUBTILASIS_SER; 1.
KM Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 697 AA; 73821 MW; B8924966C7C0714 CRC64;

Query Match 28.2%; Score 863.5; DB 2; Length 697;
Best Local Similarity 34.2%; Pred. No. 1.1e-53;
Matches 230; Conservative 98; Mismatches 215; Indels 129; Gaps 19;

QY 83 ANNKIHI-XQFNGLPEETKQKLETKGIIDYIPDAVYIYEBQDVASXXXXIEHVESV 141
DB 30 ASESDEHIIIQTAEPITABQRAELAGIDVEMQSYVDNTYLAAPPADLNRVALPVSVA 89
QY 142 BEYLPXYXIDPQLFTKGA-SKLVKAXALDTQXNKEVOL-----RGIEIXIAQXXX 190
DB 90 DYSRVRFKIPPLLRPSADTGNVRSLDHHRPDRRLERVULLHPGLEAGPELIRAVNA 149
QY 191 S-----NDVXYITAKPEYKVMNDVARGIVADVAQ 220
DB 150 AARVPEDAVAATVPGLRIITTSVGQLPEILAIDEIRIHFVRBRQLFNNVAREILNADVQL 209
QY 221 SSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALYALGRN--ANANDTNGHTIV 278
DB 210 NGTYTNGAGEVAVAVADTGFDTG-DANPHAPAFYTGVOQLYALGRAPKADDPHGHTIV 268
QY 279 AGSVLGNKXTN-----KGAPOANLVPOSIMDSXGGLGGLPSNLQTLPSQAXSAGARIH 332
DB 269 AGSVLGNKSMATMGATIGTAPALLIQLSLDPVGGIGGIPVNLNDLFQKTYDDGAYHV 328
QY 333 TNSWGA-AVNGATYDTSRNVDDVYRKNDMTILFAAGNEXPNG-----GTISAPGTA 382
DB 329 TNSWGVPEGLNTPYDASREIDFVWNNHDDVYICFAAGNDVGDGNSIGSISQSA 388
QY 383 KNATVGAETENLR-----SPGSY-----ADYNINHVAFSSRGFTYKGRIK 423
DB 389 KNCTVGAESLRKEFTSYGYTWPGDFPANPVKRDQKANNPDGVAVSSRGFTYKGRIK 448
QY 424 PDVMAFGTYILSARSSILAP-DSSFVANHSKYAVNGGSMATPIVAGNVAOLREHFVNR 482
DB 449 PDVMAFGTISILSTRNAPMONTGTSTDPLEFPDGSISMATPIVAGCAVLRRLVYNG 508

```
Qy 483 GTPKPSLLKALINGA-----ADXLGYPNGOGMGRVTLDKSLNV-----AYV 527
Db 509 LNSPAAALVKALLVGVADVLPGQINPSEAGBS-PNGNSGMGRVNLARSVLTGQPGNAGL 567
Qy 528 NESSXLSTSOXATYXF-----TATGKPLKISLWSDAPAS 563
Db 568 GEGGDFGQGEDSPFTIDPEBEVKAQAAGRNRGPAAPALPAAGVTLKTLTWSDPRGP 627
Qy 564 TTAATVLDLVDLVTTPANGIXYVGNDFKXPKXNMWGRNNVNFINKPQSGYTTIVQ 623
Db 628 -----QLNNDLDLVLVAADGSEHGN---SGTTAGFDRNNVEQVLTMTGMPGQARIYR 679
Qy 624 AYNVPVGQXFS 635
Db 680 AFRIQPPQPYA 691

RESULT 9
QANVB5_9DELTA PRELIMINARY; PRT; 1748 AA.
ID QANVB5_9DELTA PRELIMINARY; PRT; 1748 AA.
AC QANVB5;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Peptidase S8 and S53, subtilisin, kexin, sedolisin;Nepovirus coat
DE protein, N-terminal;Nepovirus coat protein, N-terminal precursor.
GN ORName=AdenDRAFT_3007;
OS Anaeromyxobacter denahogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacteriae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGR);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Istrati S., Pitsluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT denahogenans 2CP-C." to the EMBL/Genbank/DBJ databases.
RL Submitted (May-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Laft M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT denahogenans 2CP-C." to the EMBL/Genbank/DBJ databases.
RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD0100011; EAL79553.1; -; Genomic_DNA.
KW Capsid protein; Signal.
FT SIGNAL 1 23
SQ SEQUENCE 1748 AA; 177411 MW; 1B6EDB33B81DE76 CRC64;

Query Match 25.9%; Score 793; DB 2; Length 1748;
Best Local Similarity 34.9%; Pred. No. 4.5e-48;
Matches 220; Conservative 86; Mismatches 217; Indels 108; Gaps 25;

Qy 89 IXORNGFLERTKQKLEXTAKILDPYAYIVEBGVAXSXXXIXHVSVEPYLKY 148
Db 103 IVQSGFRETTHKELAAAGARLDYDLEFAVMSMDAATRAKAMLPFVGAYKAEVY 162
Qy 149 XIDPOLF-TKGASXLV-----KXALDTQXNKVEQDR 180
Db 163 KMAQQLADTSGAILEPGRSVRLVVDHDLAPLVSILVERKGMKMLHVARDAASITIL 222
Qy 181 GIEKIAQKXSNQYU-TAKREYKMDVARGIVK----ADVQASVGLYGGQIVAVA 235
Db 223 DAD-IASLANLEVLWEPALPAY-VLNDTSRWITQTVGPSPDTSISDRGLGRGQIVAI 280
Qy 236 DTGAD-----TGRDSSMHEAFRGKITALVALGRTNNANDTN-GHGTNVAGSVLGN 285
```

```
Db 281 DTGLDHDACWPRDPLGAAGEMR---KVAAYLTVC---GPDYDNLGHGHTVACTVAGD 334
Qy 286 -----GXTNKGAPQANLVPOSIMDSXGLGGLPSNLQTLFQKXSAGARIHNTSWGAA 339
Db 335 QPTITGAAAGMAGPARGVVTVDLFLGNNMFSPPADLARIPTTPYALGARHTHNSWGS 394
Qy 340 VNGAYTDSRVVDVYRKN-DMTILFPAAGNEXPKGCTISAAGTAKNALTGCAETELRPSF 398
Db 395 SN-AVDALARSADRPMHDPFLVFANGNAGPPVGSVGAATATKANVSVAGATG----- 448
Qy 399 GSYADNINHVAQFSSRGPTKDKRIKPDVMAEGTILSARSILADPSSFWAHNDSKYAYMG 458
Db 449 GLAABD---VASFSHGPADGRKTFITATGCVIVSDSCTAS-----NNCSTVAF-S 500
Qy 459 GTSNATPIVAGNVAAQLREHFNKN---RGI-----TPKPSLLKALINGA-----AADXGL 504
Db 501 GTSNATPAAAGAAALVLRQYFEGGFMPGLGSPADARSPSALVATLVNSAQNVVAGENG 560
Qy 505 G-YPNGOGMGRVTLDSLVNA-----YNBSSXLSTSOXATYXFTATGKPLKISLV 556
Db 561 GPISSTGGMGRINLSNALRPAADAAYLDVVEVAAGLETGGSFTROYVSTGAQPLKTLV 620
Qy 557 WSDAPASTTASVTLVNDLVTITAPNG-TXYVGNDFX-----XPKXNMWGRNNVNF 607
Db 621 WTDAPGSQLARSLVNDLVLVTYPGGATTYLGNVFLGSEVAGAGP-----DRLNVEQ 675
Qy 608 VFINKPQSGYTTIEVOAYNVPVGQXFSIAI 638
Db 676 VLLAAPVTGTYTVRVTGYNVVPVGPQPFALVI 706
```

```
RESULT 10
Q6MKR4_BDEBA PRELIMINARY; PRT; 711 AA.
ID Q6MKR4_BDEBA PRELIMINARY; PRT; 711 AA.
AC Q6MKR4;
DT 05-JUN-2004 (TREMBLrel. 27, Created)
DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
DE 05-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE Subtilisin-like serine protease precursor.
GN OrderedLocusNames=Bd2321;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosins A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: 11fe cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective." to the EMBL/Genbank/DBJ databases.
RL Science 303:689-692(2004).
DR EMBL; BX842652; CAB80143.1; -; Genomic_DNA.
DR HSSP; P27693; 1AH2.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILASIN.
DR PROSITE; PS00138; SUBTILASIN_SER; 1.
KW Complete proteome; Protease; Signal.
FT SIGNAL 1 8
SQ SEQUENCE 711 AA; 74914 MW; 4930050843357E5 CRC64;

Query Match 25.7%; Score 786.5; DB 2; Length 711;
Best Local Similarity 30.9%; Pred. No. 4.1e-48;
Matches 229; Conservative 110; Mismatches 246; Indels 157; Gaps 25;

Qy 9 LSVLSAALLSTVALNKPNSAGXARXFLDLPFGIGTTTDDXXGFSQXQGTGAAPLVESENV 68
```



```

Db      8 ITVTAAAVLSVA-----AQAGTVLKFNAG--AIDT-----NKLSTNSVAASMMMEKAT 55
Qy      69 KAKKGLKXKLETVPAANKKHIXPNCPPILEETQXLEXTGAKLIDYIPDYAVYIV- YEGD 127
Db      56 E-----YIVQFKKAVTEKDKAALKAQ- FEYFGYLPDDALVVRGYS 96
Qy      128 VASXXXXXIEHVESVEPYLPYXID-----PQLFTKASXLY-----KAAA 167
Db      97 LVTFKTNHNGVQAVVKTAGNYKVSNGSPAPASVFTKNNMAVLVNTFKSHEVEVIAKTEK 156
Qy      168 LPTKQXKEVQ-----LRGIEXI-----AQXXSNDVXYIT 198
Db      157 MDVKVAVQVVDGKHITALLIPRGVLPAVALVTVENHQPAPAVISLIRPMDEDLADQVSAT 216
Qy      199 APEPYVMNDVARG--IVKADVAQSSYGLYGQOIVAVADTGLDTRGNDSSMHEAFRGKI 256
Db      217 AAGDYSDDLTVGDESGLTMNFDAAW-AMGYAGKQQTVMADTGLDSC-NTGALHQDPAGGV 274
Qy      257 TALVYALGRNNA-NDPNGHGVHAGSVLGNKXTNKM-----APQANLVFQSIIMDSXGGLG 311
Db      275 IGGYFPGLSKMSKSDPMGHTHVASVWGKGTASKGLKGAYEAMNVVAEGMWSPMKQUL 334
Qy      312 GLPSNIQTLFQXKXASAGARIHTNSMGA- VNGAYTTDSRVDDYVRKN-DMTILPAAGNH 369
Db      335 SVPSKLGDLFEKAFADGARIHTNSKSGARTFGAYDPAVQVDEMSYANPDLILPAAGNS 394
Qy      370 XP-----NGGTISAPGTAKNATVGAATENI-----RPSFG 399
Db      395 GADKNKDKGIDNSNMSAPGTAKNATVVGASENVTKSGGIQVPISKRAADWEPSEPIYS 454
Qy      400 SY-ADNINVAQSSSGPTKDKGIRKPDVWAPGTXYLSARSSSLAPDSSFPAMHDSKAYNG 458
Db      455 STISDGNGLAMPSSSGPTTGDGKPDYVAPGTNVLSVPSQEKDAPLWGANVKKDYVWSG 514
Qy      459 GTSMATPIVAGNVAQLREHFVNKRG-TRKPSILKALJAGADKGLG-----506
Db      515 GTSMATPLAAGAAALARQVLEKGMKNSALMKRTMLHTAVDMTPQGPGEIGARQGE 574
Qy      507 -----PNCQMGKRVTLDKSLVA-----YVNBSSKXISTOKATYKFTATAGKPLKISLVW 557
Db      575 ILTRPNSDEGKRVADVANIANGATQGVDRNKGVAQGAQAVSYEBTLNAPGSLVNLVW 634
Qy      558 SDAPSTTASVTLVNDLIVITAPNGTXVVGDPFKPXXKXNDGRNVENVPIKXQSGT 617
Db      635 TDAPSANAAQALVNDLIVLTPNGQTLISMDHI-----NNLEMKRSGLPACT 684
Qy      618 YTIKVAQVVPV---GPOKFSL 636
Db      685 YKLVTKGPRVPOGKNGAKAYAL 706

RESULT 11
Q747P6 GEOSL PRELIMINARY; PRT; 2030 AA.
AC 0747P6-
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DB Fibronectin type III domain protein.
GN OrderedLocustName=GSU3219;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,

```

```

RA Davidson T.M., Zafer N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.B., Loyley D.R., Frazer C.M.,
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments."
RL Science 302:1967-1969(2003).
DR EMBL; AE017180; JAR36610.1; -; Genomic_DNA.
DR HSSP; P27693; IAH2.
DR TIGR; GSU3219; -.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR011635; APHP.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF07705; CARD; 8.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
KW Complete proteome.
SQ
SEQUENCE 2030 AA; 207213 MW; BD7ADC27DD141B01 CRC64;

Query Match      20.9%; Score 638; DB 2; Length 2030;
Best local similarity 29.4%; Pred. No. 8 6e-37;
Matches 188; Conservative 93; Mismatches 243; Indels 116; Gaps 19;

Qy      83 ANKHLHXQFNPPILEETQXLEXTGAKLIDYIPDYAVYVEYEGDVASXXXXXIEHVESVE 142
Db      54 AIRKRLVQFNGPVRPEGRQLEALGCRIGDYMPTFAFVALMDDKAAKQVALLSPVEDIT 113
Qy      143 PLYPKYXIDPOLFTKASLYKAXALD-TRKQNKVEQLAGI-----182
Db      114 RFAF-----ADTLVGTARKDLTAAPTSIRIKRVLRVDDPADRAAVIAATLNG 161
Qy      183 -----EXIAQXXSNDVXYITAKPEYKVMNDVARGIVKADVAQS---221
Db      162 NGRILNVGARTITVEPBELLAPLAQOESTAMIGVGLRLNSDAAWVQVTRNEVDNRIT 221
Qy      222 -SYGLYGQOIVAVADTGLD-----TGRNDSMHEAFRGKITALVYALGRTNANDT 271
Db      222 WEKGITGAGQIVIGIDSGVDYDMPFADPENGALPGGRHKIYG-----YDATLGDHDA 276
Qy      272 NCHGTHVASVVGN---GXTKNGMAPQANLVFQSIIMDSVSGGLGSLPSNIQTLFQXKXSG 328
Db      277 DGHGTHIGETICGDBGPGMGNGIAPGARIHVQDVGDTGLTG-SLELETYLKRAYDSG 335
Qy      329 ARHTNSMGAVNGAYTTDSRVDDYV-RKNDMTILPAAGNEXPNCGTISAPGTAKNAT 387
Db      336 ARIFGSMGVD-SGNYDALLAALDDPSWRHKDPLAVFANGCGPABQIATSPALAKNAT 394
Qy      388 VGATENLRPSPFGSYADNINVAQFSSRGPTKDKGIRKPDVWAPGTXYLSARSSSLAPDSSFW 447
Db      395 VVATGN-----GTDAAI---VSAESSVGOAPDGRANPSVGARQGVASARS---DGLLG 442
Qy      448 ANHDSKAYVMGTSNATPIVAGNVAQLREHF-----YKNGITPKPSLLKALIA 497
Db      443 SGNSTMA-MSGTSVAALVTSGAALIRQYFDGFPFGSPATVTKLPASALLKAVLVN 501
Qy      498 GAA-----DXGLGYVNGNMGKRVTLDKSLV-----NVAAYVNBSSKXISTOKATYKFT 545
Db      502 SAEALLSDPDGSCSKGKGMGRPKLINTLFFNGSHSLVAVDGGTGLTGVQWQRLYS 561
Qy      546 TAGKPLKISLVSDAPASTTASVTLVNDLIVITAPNGTXVVGNDP-----XEPKXNW 599
Db      562 PGRRLKITLAWTADPAAPAGATSPLTNDLNLVVAAPDGTTLGNLNSHGDYESRTGSG 621
Qy      600 DGRNV-ENVPIKXQSGTITIEVQAVVPVPGQKFSLA 638
Db      622 SDRVNVEQVYIKRPVAGTYLVKVIKVASIPVGPOPFALVM 661

RESULT 12
Q4H0U5_GTBZ

```

ID 04HUY5 GIBZE PRELIMINARY; PRT; 1088 AA.
AC 04HUY5-
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG11223.1;
OS Gibberella zeae PH-1
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocryomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TextID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barua N., Bastien V., Bloom T., Boguslavsky L.,
RA Boughalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Chepel Y., Collymore A., Cook A., Cooke P., Corum B., DeRellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardina S., Gerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-Y., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mäkelä E., McCarthy M., Meidrim J., Meneus L.,
RA Mitrova T., Miedge V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Nordu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachtupka A., Ramsay U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schuer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talama J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataran V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACW01000460; EAA75433.1; -! Genomic_DNA.
CC Hypothetical protein.
SQ SEQUENCE 1088 AA; 119629 MW; E3B3CB94C07F542 CRC64;
Query Match 19.4%; Score 594; DB 2; Length 1088;
Best Local Similarity 31.8%; Pred. No. 5,8e-34;
Matches 197; Conservative 98; Mismatches 231; Indels 94; Gaps 29;
QY 96 ILEETKQLEKTKGAKILDIYIPYAYIVE-YEGDVXSKXXTEHVSVEPYLPXYXIDPOL 154
DB 476 ILITQK---LTGSPYLAEKSTISLKMWDRELDRTYRDCILHEPNTYAEKLAPEI 532
QY 155 FTGKASXLVKKALDPTQKXNKEVQGLGEXIAQXXXNDVYXTAKPEKYMNVDVARGIV 214
DB 533 AEKAGVDITK-LAVSPVRIRLTVHDDKLEALAKLDSIRIEBV-RPD-EVTLNDLARRTL 588
QY 215 KADVAQSSYGLYGQGIYAVADTGLDTRGNSSM---HEAFRGKITLALYALGRNNAND 270
DB 589 NANTIALSTSYEGNGQKCVADTPGQKMDDEMIILVHPFNGVEHLEALM-LGDSKD 647
QY 271 TNGGTHVAGSVLGNKXTN-----KGNAPQANLIVPQSI-----MDSXGLGLPSNL-QT 319
DB 648 TAGGTHVCASICGGLYKNGDIRKRGVAPGATLMVQSIAGVSRPNKGALEVPMDLQ 707
QY 320 LFSQAKSAGARIHTNSGA---AVNG-AYTTSRNVDDY-RKNDMTILPAAGNEKXNG 373
DB 708 LFSNPYKGYRIHSHSGWKVMDAKTGQLEGQAMDDIKFYIDHDFVVLVAAGNNAKA 767
QY 374 GT-----ISAPGTAKNAITVGAATENRPSFGSYADN-----INHAQSSSGPTD- 419
DB 768 KSKSNHIGAAQSAFNCITVGAITGTRPNNDYGFQDNEVGAQPMTRINDTAKFSSRGPTKG 827

QY 420 -----GRKPDVMAPGTXILSARS-SLAPDS-----SFWANHDSKAYMGTSMATP 465
DB 828 RDINGNEYAGRIKPDVYAPVAILLSAASRAAAKQSRNVMYGRGDDMTWMTSGTMSSTP 887
QY 466 IVAGNVQLRHFVYNNRITPKPSILKALJAGAAD-----XGIGYNGNGMGWRYLTD 519
DB 888 LVAGCVALLRALKHENGKPSALLIKYLVNGAVNPSQGLGIGY-DYQGFGRVID 946
QY 520 KSLNV---AYNESSKLISTQ-----KATYFTAT-----AGK-PLKISLWSD 559
DB 947 SSISSVVKLSFVDDGKLFEDYQFVAPLRQVPEERERWTSLSIIPACRNRLLTVTLAYPD 1006
QY 560 APASTTAVTLVNDLIVTAPNGTXVYGVNPFXXPXKXNMGRNVVFINXPGSGYVT 619
DB 1007 KPAQ---SGLMQNDINLVLVS-GAHERHGNMKGP---GYDHTNVKTIWENVPGETFK 1059
QY 620 IEVOAY-NVPY-GPOXFSIA 637
DB 1060 IVASIMWNIDVKAPTSFAVA 1079

RESULT 13
AC 054M84 D1CIDI
ID 054M84 D1CIDI PRELIMINARY; PRT; 1741 AA.
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE ABC transporter B family protein.
GN Name=tagC; ORFNames=DDB0191192;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellidae; Dictyostellium.
OX NCBI_TextID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sungug B., Bertrian M., Song J., Olsen R., Szafrenski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivery F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kethornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hanser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loughsagh H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simonde M., Splegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kaya R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrett B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum."
RL Nature 0-0-0(2005).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAF01000133; EAL64353.1; -! Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR0003439; ABC_transp_like.
DR InterPro; IPR000209; Pept_S8_S53.

FT COMBRIAS 1353 1357 Poly-Asn.
FT COMBRIAS 1358 1364 Poly-Asp.
FT COMBRIAS 1361 1386 Poly-Asn.
FT COMBRIAS 1707 1729 Poly-Asn.
FT ACT SITE 325 325 Charge relay system (By similarity).
FT ACT SITE 372 372 Charge relay system (By similarity).
FT ACT SITE 397 397 Charge relay system (By similarity).
FT CARBOHYD 330 637 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 536 536 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 547 547 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 614 614 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 689 689 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 741 741 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 776 776 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 832 832 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 887 887 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1251 1251 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1385 1385 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1386 1386 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1454 1454 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1704 1704 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1743 AA, 194146 MW, 12DB363E2F729839 CRC64;

Query Match 16.9%; Score 518.5; DB 1; Length 1743;
Best Local Similarity 24.2%; Pred. No. 3e-28; Indels 281; Gaps 33;
Matches 194; Conservative 114; Mismatches 213;

QY 88 HIXQNGPILEETKQKL-----EXTGAKIIDYIPDYAYIV---EYEGDVXK 131
DB 115 YIVGPKKINDETRQLKEPLIGTDIVLDEQRYGSHIYHYPHBSFLVLMQEOGVLLSS 174
QY 132 XXXIEHVESVPEY-----LPXYIIDPOLFTKASXLVK--AKALDT--KQXN 174
DB 175 KEWYWIQEPSPNKHINYNKSIQLEPYII-----LSDSTNSLIQRMENLINSILKSYN 230
QY 175 KEVQLRGI-----EXIAQXXSXNDVYITAKPEYK 204
DB 231 SKVKLTLLNQKKLSIYVCSNDESPSCSLINSEKLVYQWISSEQSEN--FIERSEKQ 287
QY 205 VMDNVARGIV-----KADVAQSSYGLYGGQIVAAVADTLDJGR-----NDS----- 246
DB 288 TANLSPRVVPGTQDTLVNNDKVDLP-----LRKGQGLISTADTGLDSHCFPSISKPI 342
QY 247 ---SMHEAFRKITALVALGRNNANDNGTGHVAGSVLG-----NGXTNKGAPQA 296
DB 343 PLNSVNLNHR--KVYVYITTSDDSDKDKYCGHICGSAAGPBDSPVNISSFGIATDA 401
QY 297 NLVQSIWDSXGLGGL--PSNLQTLFQQAASAGARHITNSMGA---AVNGAYTTDSRN 350
DB 402 KIAF--FDLASGSSSLTPPSDLKQLYPLDYAGARVHCDMSGVSVSGYTSYSSDTAS 458
QY 351 VDDYVRKN-DMTILFAAGNEXPNQGTIS--APGTAKNIYVQATENLR-----PSRFS 400
DB 459 IDDLFLFHPDITILRAAGN---NEQYLSLTLQSTAKVITVGAQHTIHENVLTDGPNIN 515
QY 401 YADNI----- 405
DB 516 YQSSVFNQELICDFSRVYCYTTAQCCLASNAATTGLASCPTLLRKSVIDAANTQPLLY 575
QY 406 --NHVAQSSRGPTKDGRIKPDVMAFGTXIISABSLA-----PDSSFMANHDSKVYM 457
DB 576 NENNI CSFSSKGFTHDGMKPALVAPGYITTSARNGNANTTQCCDGL--PVTNALA-I 633
QY 458 GGTSMATPIVAGNAQLREH-----FVKNGITPKPSILKAALIAA----- 499
DB 634 SGTMATSPFAAATITILKQYLVDGYTPGSGIVESKCKOPTGSLIKALMINNAQLNGTFQ 693
QY 500 --ADXLGAYPNGN-----QGWGRVTLDSLVNAYVNESG----- 531
DB 694 LITSSSIYPSNQVFNENFAGASLVQGMGAIHSMNLHVNNNNNSNNNKTSIDGTIKFDGI 753
QY 532 -----XLTSTQKATYXT-----ATAGKPRK-----ISLVMSDAPASTT 565

DB 754 GGLDLRLVKNQWKEBSLSTQNTSYCTTYPSSSSSSNSGNNIRRVAVLTWTPPSTAG 813
QY 566 ASVLVNDLPLVI-----TAPNGTXVYVNDPFXKXKXNMDGRNVENVFINXP 613
DB 814 AKFPLVNNLDTMTLYRNDNGSTIRYSNQGSSFLG-----LAPQDTLVNVEGIVHNPT 867
QY 614 QSGTYTTEVQAYNVNVPVQKFS 635
DB 868 EPMYTRFVAVGTNVMPGPNPFS 889

RESULT 15
ID Q8T9W1_DICDI PRELIMINARY; PRT; 1825 AA.
AC Q8T9W1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Serine protease/ABC transporter tagB (ABC transporter B family protein).
GN Name=tagB; ORFNames=DD80191427;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Eukaryota; Dictyosteliales; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Anjard C.; Loomis W.F.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L.; Pachbat J.A.; Gloeckner G.; Rajandream M.-A.;
RA Sugarcang R.; Berriman M.; Song J.; Olsen R.; Szafrenski K.; Xu Q.;
RA Tunggal B.; Kummerfeld S.; Madera M.; Konfortov B.A.; Rivero F.;
RA Bankier A.T.; Lehmann R.; Hamlin N.; Davies R.; Gaudet P.; Fey P.;
RA Plichter K.; Chen G.; Saunders D.; Sodergren E.; Davis P.;
RA Kechroian A.; Nie X.; Hall N.; Anjard C.; Hemphill L.; Bason N.;
RA Fatholmou A.; Desany B.; Just E.; Morio T.; Rost R.; Churcher C.;
RA Cooper J.; Haydock S.; van Driessche N.; Cronin A.; Goodhead I.;
RA Murthy D.; Mourier T.; Pain A.; Lu M.; Harper D.; Lindsay R.;
RA Hauser H.; James K.; Quiles M.; Mohan M.B.; Saito T.; Buchrieser C.;
RA Waidrop A.; Felder M.; Thangavelu M.; Johnson D.; Knights A.;
RA Louisge H.; Mungall K.; Oliver K.; Price C.; Quail M.A.;
RA Urushihara H.; Hernandez J.; Rabinowitsch E.; Steffen D.; Sanders M.;
RA Ma J.; Kohara Y.; Sharp S.; Simmonds M.; Splegler S.; Rivey A.;
RA Shulsky G.; Schleicher M.; Woodward J.; Munkler T.; Tanaka Y.;
RA Chisholm R.L.; Gibbs R.; Loomis W.F.; Platzner M.; Kay R.R.;
RA Williams J.; Dear P.H.; Noegel A.A.; Barrett B.; Kuapa A.;
RL "The genome of the social amoeba Dictyostelium discoideum";
EMBL; AF466309; AAL74253.1; -; Genomic DNA.
EMBL; AAT01000133; EAL64354.1; -; Genomic DNA.
HSP; P08716; IMT0.
DR Dictybase; DD80191427; tagB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0046262; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008283; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR000209; Pept_S8_S53.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran; 1.

This Page Blank (uspto)


```
Db 241 TGRDSSMHEAFRGKITLALYALGRNNANDTNGHGTTHAGSVLNGSTNKGMAPOANLVF 300
Qy 301 QSIMDSXGGLGGLPSNLQTLFSGAASAGARJHTNSWGAAVNGAYTTDSRNVDVYRKNDM 360
Db 301 QSIMDSXGGLGGLPSNLQTLFSGAASAGARJHTNSWGAAVNGAYTTDSRNVDVYRKNDM 360
Qy 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQSSRPTXDG 420
Db 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQSSRPTXDG 420
Qy 421 RIKPDVMAPGTIIISARSSILAPDSSFMANHDSKYAYMGSTMATPIVAGNVAQLREHFVK 480
Db 421 RIKPDVMAPGTIIISARSSILAPDSSFMANHDSKYAYMGSTMATPIVAGNVAQLREHFVK 480
Qy 481 NRGITPKPSLLKAAIAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSSLSTSOKAT 540
Db 481 NRGITPKPSLLKAAIAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSSLSTSOKAT 540
Qy 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYXGNDPXXKXNWD 600
Db 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYXGNDPXXKXNWD 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGQPSLAIVN 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGQPSLAIVN 640
```

RESULT 2
US-09-920-954-6
Sequence 6, Application US/09920954
Patent No. 6759228

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-6

Query Match 99.0%; Score 3029; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 3e-258;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy 1 MRXKKQVFLSYLSAAAILSTVALKNPSAGAKXFDLPFGKIQTITDDXXGFSKQXGTGAAA 60
Db 1 MRXKKQVFLSYLSAAAILSTVALKNPSAGAKXFDLPFGKIQTITDDXXGFSKQXGTGAAA 60
Qy 61 FLVSENVKLLKGLKKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
Db 61 FLVSENVKLLKGLKKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
Qy 121 IVEYEGDVKXSKXXIETHESVEPYLPYXIIDPOLFTKGASLVKAXALDTKQXNKEVOLR 180
Db 121 IVEYEGDVKXSKXXIETHESVEPYLPYXIIDPOLFTKGASLVKAXALDTKQXNKEVOLR 180
```

```
Db 121 IVEYEGDVKXSKXXIETHESVEPYLPYXIIDPOLFTKGASLVKAXALDTKQXNKEVOLR 180
Qy 181 GIEKIAQXXXNDYXYITAKPEYRVMDVANGIYKADVAQSSYGLYGGQIIVAADTGLD 240
Db 181 GIEKIAQFAISNDVLYITAKPEYKVMNDVAGIYKADVAQSSYGLYGGQIIVAADTGLD 240
Qy 241 TGRDSSMHEAFRGKITLALYALGRNNANDTNGHGTTHAGSVLNGSTNKGMAPOANLVF 300
Db 241 TGRDSSMHEAFRGKITLALYALGRNNANDTNGHGTTHAGSVLNGSTNKGMAPOANLVF 300
Qy 301 QSIMDSXGGLGGLPSNLQTLFSGAASAGARJHTNSWGAAVNGAYTTDSRNVDVYRKNDM 360
Db 301 QSIMDSXGGLGGLPSNLQTLFSGAASAGARJHTNSWGAAVNGAYTTDSRNVDVYRKNDM 360
Qy 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQSSRPTXDG 420
Db 361 TILPAAGNEXPNGGTISAPGTAKNAITVGATENLRPSFGSYADININHAQSSRPTXDG 420
Qy 421 RIKPDVMAPGTIIISARSSILAPDSSFMANHDSKYAYMGSTMATPIVAGNVAQLREHFVK 480
Db 421 RIKPDVMAPGTIIISARSSILAPDSSFMANHDSKYAYMGSTMATPIVAGNVAQLREHFVK 480
Qy 481 NRGITPKPSLLKAAIAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSSLSTSOKAT 540
Db 481 NRGITPKPSLLKAAIAGADIGLGYPNGNGMGRVTLDKSLNVAAYNESSSLSTSOKAT 540
Qy 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYXGNDPXXKXNWD 600
Db 541 YXFTATAGKPLKISLWSDAPASTTASVTLVNDLVLITAPNGTYXGNDPXXKXNWD 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGQPSLAIVN 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGQPSLAIVN 640
```

RESULT 3
US-09-509-814A-8

Sequence 8, Application US/09509814A
Patent No. 6376227

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-509-814A-8

Query Match 99.0%; Score 3028; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 3.7e-258;
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
Qy 1 MRXKKQVFLSYLSAAAILSTVALKNPSAGAKXFDLPFGKIQTITDDXXGFSKQXGTGAAA 60
Db 1 MRXKKQVFLSYLSAAAILSTVALKNPSAGAKXFDLPFGKIQTITDDXXGFSKQXGTGAAA 60
Qy 61 FLVSENVKLLKGLKKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
Db 61 FLVSENVKLLKGLKKLETPANNKLIHQFNPILEBTQXLEXTGAKILDIYIPDAY 120
```



```
|||||
Db 61 FLVSESNVLPKGLQKQLFTVPANNKLTIIQFNGPILFETTKQLETKGAKIIDYIPDVAY 120
Qy 121 IVEBEGDVYKXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVKAYALDTPKXNEVOQR 180
Db 121 IVEBEGDVYKASSTIEHVESVEPYLPYRIDPOLFTKGASELVKAVALLDTPKXNEVOQR 180
Qy 181 GIEIXAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQOQIVAVADTGLD 240
Db 181 GIEIXAQFPAISNDVLYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQOQIVAVADTGLD 240
Qy 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTKGMAPQANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTKGMAPQANLVF 300
Qy 301 OSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Db 301 OSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Qy 361 TILPAAGNEXPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Db 361 TILPAAGNEXPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Qy 421 RIKPDVMAFGTYILSARSSLPDSSFMANHDSKYAVMGTSMAATPIVAGNVAQLEHFPK 480
Db 421 RIKPDVMAFGTYILSARSSLPDSSFMANHDSKYAVMGTSMAATPIVAGNVAQLEHFPK 480
Qy 481 NRGITPKPSLLKALAIAGAADYGLGYPNGNQGKGRVTLDKSLNVAVYNESSXLSTSQKAT 540
Db 481 NRGITPKPSLLKALAIAGAADYGLGYPNGNQGKGRVTLDKSLNVAVYNESSXLSTSQKAT 540
Qy 541 YFPTAGKPLKISLWMSDAPASTTASVTLVNDLVLVITAPNGTXYVGNDFXXPXXKND 600
Db 541 YFPTAGKPLKISLWMSDAPASTTASVTLVNDLVLVITAPNGTXYVGNDFXXPXXKND 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQFSLATVN 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQFSLATVN 640
```

RESULT 4
US-09-920-954-8
Sequence 8, Application US/09920954

GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920, 954
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509, 814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-8

Query Match 99.0%; Score 3028; DB 2; Length 640;
Best Local Similarity 93.3%; Pred. No. 3.7e-258;

```
Matches 597; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
Qy 1 MEXKKKKVFLSVSAALIISTVALKNPSAGKAPXFLDFFGIOTTTDXKFSQXQTGA 60
Db 1 MEXKKKKVFLSVSAALIISTVALKNPSAGKAPXFLDFFGIOTTTDXKFSQXQTGA 60
Qy 61 FLVSESNVLPKGLQKQLFTVPANNKLTIIQFNGPILFETTKQLETKGAKIIDYIPDVAY 120
Db 61 FLVSESNVLPKGLQKQLFTVPANNKLTIIQFNGPILFETTKQLETKGAKIIDYIPDVAY 120
Qy 121 IVEBEGDVYKXXXIIEHVESVEPYLPXYXIDPOLFTKGASXLVKAYALDTPKXNEVOQR 180
Db 121 IVEBEGDVYKASSTIEHVESVEPYLPYRIDPOLFTKGASELVKAVALLDTPKXNEVOQR 180
Qy 181 GIEIXAQXXXSNDVXYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQOQIVAVADTGLD 240
Db 181 GIEIXAQFPAISNDVLYITAKPEYKVMNDVARGIVKADVAAQSSYGLYGQOQIVAVADTGLD 240
Qy 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTKGMAPQANLVF 300
Db 241 TGRNDSMHEAFRGKITLALYALGRTNNANDTNGHGVAGSVLGNKXTKGMAPQANLVF 300
Qy 301 OSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Db 301 OSIMDSXGGLGGLPSNLQTLFQOAXSAGARIHTNSMGAAVNGAYTTDSHNDVDDYRKNDM 360
Qy 361 TILPAAGNEXPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Db 361 TILPAAGNEXPNCGTISAPGTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTKDG 420
Qy 421 RIKPDVMAFGTYILSARSSLPDSSFMANHDSKYAVMGTSMAATPIVAGNVAQLEHFPK 480
Db 421 RIKPDVMAFGTYILSARSSLPDSSFMANHDSKYAVMGTSMAATPIVAGNVAQLEHFPK 480
Qy 481 NRGITPKPSLLKALAIAGAADYGLGYPNGNQGKGRVTLDKSLNVAVYNESSXLSTSQKAT 540
Db 481 NRGITPKPSLLKALAIAGAADYGLGYPNGNQGKGRVTLDKSLNVAVYNESSXLSTSQKAT 540
Qy 541 YFPTAGKPLKISLWMSDAPASTTASVTLVNDLVLVITAPNGTXYVGNDFXXPXXKND 600
Db 541 YFPTAGKPLKISLWMSDAPASTTASVTLVNDLVLVITAPNGTXYVGNDFXXPXXKND 600
Qy 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQFSLATVN 640
Db 601 GRNNVENVFINKPQSGTYTIEVOAYNVPGPQFSLATVN 640
```

RESULT 5
US-09-509-814A-4

Sequence 4, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 639
TYPE: PRT

ORGANISM: Bacillus sp.
US-09-509-814A-4

Query Match 98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1.9e-257;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

Qy 4 KKKVPLSLTSAALISTVALKNPSAGARXPDLPFGKIQTITDXXGSPKQOTGAAPLV 63
Db 3 KKKVPLSLTSAALISTVALKNPSAGARXPDLPFGKIQTITDXXGSPKQOTGAAPLV 62
Qy 64 ESENVKLKGLKGLKLETPANNKLIHQFNGPILEETKQLEXTGAKILDIYIPYAYIVE 123
Db 63 ESENVKLKGLKGLKLETPANNKLIHQFNGPILEETKQLEXTGAKILDIYIPYAYIVE 122
Qy 124 YEGDVXSXXXIIEHVESVEPYLPYXXIDPOLFTKGASLTVKXALDTKQNKVEVLRGIE 183
Db 123 YEGDVQSXRSIEHVESVEPYLPYXXIDPOLFTKGASLTVKXALDTKQNKVEVLRGIE 182
Qy 184 XIAQXXXSNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGGQIIVAADTGLDTR 243
Db 183 EIAQVANSNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGGQIIVAADTGLDTR 242
Qy 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFQSI 303
Db 243 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFQSI 302
Qy 304 MDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 363
Db 303 MDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 362
Qy 364 PAAGNEXPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSRGPTKDGRIK 423
Db 363 PAAGNERPNGGTISAPGTAKAALITVGATENLRPSFGSTADININHAQFSSRGPTKDGRIK 422
Qy 424 PDVNAFGTYILSARSSSLAPDSSFMANHDSKXYAVMGTSMTPTPIVAGNVAQLEHFFVKNRG 483
Db 423 PDVNAFGTYILSARSSSLAPDSSFMANHDSKXYAVMGTSMTPTPIVAGNVAQLEHFFVKNRG 482
Qy 484 ITPKPSLLKALILAGAADXGIGYPNGQGWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 543
Db 483 ITPKPSLLKALILAGAADXGIGYPNGQGWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 542
Qy 544 TATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXXPXXNWDGRN 603
Db 543 TATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTRYVGNDFSAFPDNNWDGRN 602
Qy 604 NVENVFINKPQSGTYTIEVQAYNVPGPQXSLAIYN 640
Db 603 NVENVFINKPQSGTYTIEVQAYNVPGPQXSLAIYN 639

```

RESULT 6
US-09-920-954-4
Sequence 4, Application US/09920954
Patent No. 6759228
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US/09/920,954
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-920-954-4

Query Match 98.7%; Score 3020; DB 2; Length 639;
Best Local Similarity 93.4%; Pred. No. 1.9e-257;
Matches 595; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

Qy 4 KKKVPLSLTSAALISTVALKNPSAGARXPDLPFGKIQTITDXXGSPKQOTGAAPLV 63
Db 3 KKKVPLSLTSAALISTVALKNPSAGARXPDLPFGKIQTITDXXGSPKQOTGAAPLV 62
Qy 64 ESENVKLKGLKGLKLETPANNKLIHQFNGPILEETKQLEXTGAKILDIYIPYAYIVE 123
Db 63 ESENVKLKGLKGLKLETPANNKLIHQFNGPILEETKQLEXTGAKILDIYIPYAYIVE 122
Qy 124 YEGDVXSXXXIIEHVESVEPYLPYXXIDPOLFTKGASLTVKXALDTKQNKVEVLRGIE 183
Db 123 YEGDVQSXRSIEHVESVEPYLPYXXIDPOLFTKGASLTVKXALDTKQNKVEVLRGIE 182
Qy 184 XIAQXXXSNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGGQIIVAADTGLDTR 243
Db 183 EIAQVANSNDVXYITTAPEYKVMNDVARGIVKADVAOSSYGLYGGQIIVAADTGLDTR 242
Qy 244 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFQSI 303
Db 243 NDSMHEAFRGKITLALYALGRTNANDTNGHGTIVAGSVLGNKXNKGMAPQANLVFQSI 302
Qy 304 MDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 363
Db 303 MDSXGGLGGLPSNLQTLFSGQXSAGARIHTNSWGAAVNGAYTTDSRVDDYVRKNDMTIL 362
Qy 364 PAAGNEXPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSRGPTKDGRIK 423
Db 363 PAAGNERPNGGTISAPGTAKAALITVGATENLRPSFGSYADININHAQFSSRGPTKDGRIK 422
Qy 424 PDVNAFGTYILSARSSSLAPDSSFMANHDSKXYAVMGTSMTPTPIVAGNVAQLEHFFVKNRG 483
Db 423 PDVNAFGTYILSARSSSLAPDSSFMANHDSKXYAVMGTSMTPTPIVAGNVAQLEHFFVKNRG 482
Qy 484 ITPKPSLLKALILAGAADXGIGYPNGQGWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 543
Db 483 ITPKPSLLKALILAGAADXGIGYPNGQGWGRVTLDKSLNVAAYNNESSSLSTSOKATYTF 542
Qy 544 TATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTYVGNDFXXPXXNWDGRN 603
Db 543 TATAGKPLKISLWSDAPASTTASVTLVNDLDTVITAPNGTRYVGNDFSAFPDNNWDGRN 602
Qy 604 NVENVFINKPQSGTYTIEVQAYNVPGPQXSLAIYN 640
Db 603 NVENVFINKPQSGTYTIEVQAYNVPGPQXSLAIYN 639

```

RESULT 7
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
TITLE OF INVENTION: ALKALINE PROTEASE

```

FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/045528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRF
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)

```

```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

```

Query Match

98.6%, Score 3016, DB 2, Length 640;

```
Best Local Similarity 100.0%; Pred. No. 4.2e-257;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRKKKKVFLSVLSAAALISTVALXNPSSAGKAXFPDLDFKGIOTTTDXXGFSKQXGTGAAA 60
DB 1 MRKKKKVFLSVLSAAALISTVALXNPSSAGKAXFPDLDFKGIOTTTDXXGFSKQXGTGAAA 60
QY 61 FLVSEENVKLAKGLKKLETTVPANNKLAHXQNGPLLEETKQXLEXTGAKIIDYIPDAY 120
DB 61 FLVSEENVKLAKGLKKLETTVPANNKLAHXQNGPLLEETKQXLEXTGAKIIDYIPDAY 120
QY 121 IVEYEGDVAXXXXXIEHVESVEPYLPHYXIDIDQLFTKGASVLVKAALDTKQXNREVOLR 180
DB 121 IVEYEGDVAXXXXXIEHVESVEPYLPHYXIDIDQLFTKGASVLVKAALDTKQXNREVOLR 180
QY 181 GLEXIAQXXXXNDVYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQGIIVAVADTGLD 240
DB 181 GLEXIAQXXXXNDVYITAKPEYKVMNDVARGIVRADVAQSSYGLYGQGIIVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALGRNNANDTNGHGHVAGSVLGNGXTNKGMAPOANLVF 300
DB 241 TGRNDSMHEAFRGKITLALGRNNANDTNGHGHVAGSVLGNGXTNKGMAPOANLVF 300
QY 301 OSIMDSXGGLGGLPSNLQTLFSGQAXSAGARIHTNSGAAVNGAYTTDSRNDVDYRKNDM 360
DB 301 OSIMDSXGGLGGLPSNLQTLFSGQAXSAGARIHTNSGAAVNGAYTTDSRNDVDYRKNDM 360
QY 361 TILFAAGNEXPNGGTISABGTAKNAITVGATENLRPSFGSYADININHAQSSRGPTKDG 420
DB 361 TILFAAGNEXPNGGTISABGTAKNAITVGATENLRPSFGSYADININHAQSSRGPTKDG 420
QY 421 RIKPVTMAPGTITILARSLLAPDSFMANHDSKAYMGTSMAATPIVAGNVQLBEHPK 480
DB 421 RIKPVTMAPGTITILARSLLAPDSFMANHDSKAYMGTSMAATPIVAGNVQLBEHPK 480
QY 481 NRGITPKPSILKAALIAGAADKGLGYPNNGQMGRTVLDKSLNVVYVNESSLSTSOXAT 540
DB 481 NRGITPKPSILKAALIAGAADKGLGYPNNGQMGRTVLDKSLNVVYVNESSLSTSOXAT 540
QY 541 YXFTATAGPKLKLISLVMSDAPASTASTAVTLVNDLVLITAPNGTYVGNDFXXPXXKWD 600
DB 541 YXFTATAGPKLKLISLVMSDAPASTASTAVTLVNDLVLITAPNGTYVGNDFXXPXXKWD 600
QY 601 GRNNVENVPINXPOSGTITIEVQATNVPGPOXFSLATV 640
DB 601 GRNNVENVPINXPOSGTITIEVQATNVPGPOXFSLATV 640

RESULT 8
US-09-920-954-2
/ Sequence 2, Application US/09920954
/ Patent No. 6759228
/ GENERAL INFORMATION:
/ APPLICANT: TAKAIWA, MIKIO
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAKKI, KATSUISHA
/ APPLICANT: KUBOTA, HIROMI
/ APPLICANT: HITOMI, JUN
/ APPLICANT: KAGEYAMA, YASUSHI
/ APPLICANT: SHIKATA, SHITSUM
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 0327-0832-0PCT
/ CURRENT APPLICATION NUMBER: US/09/930,954
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: PCT/J998/04528
/ PRIOR FILING DATE: 1998-10-07
/ PRIOR APPLICATION NUMBER: JP 9-274570
/ PRIOR FILING DATE: 1997-06-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.0
```

```
SEQ ID NO 2
/ LENGTH: 640
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
/ NAME/KEY: misc_feature
/ LOCATION: (3)..(3)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (24)..(24)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (30)..(30)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (33)..(33)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (47)..(47)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (48)..(48)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (54)..(54)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (71)..(71)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (75)..(75)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (90)..(90)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (103)..(103)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (106)..(106)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (129)..(129)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (131)..(131)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (132)..(132)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (133)..(133)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (134)..(134)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (147)..(147)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (149)..(149)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (161)..(161)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (166)..(166)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (173)..(173)
/ OTHER INFORMATION: Xaa is any amino acid
/ NAME/KEY: misc_feature
/ LOCATION: (184)..(184)
```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-2

Query Match 98.6%; Score 3016; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 4.2e-257;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEXKKVPISTVSAALISTVALKNPSAKAXKPDLDPRGIGTITDDXKFSKXQTGA 60
DB 1 MEXKKVPISTVSAALISTVALKNPSAKAXKPDLDPRGIGTITDDXKFSKXQTGA 60
QY 61 FLVSEENVTLKKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 120

DB 61 FLVSEENVTLKKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGLKGL 120
QY 121 IVEYEGDVAXXXXXXIEHVESVPEPLPXYYIDPOLFTKGASXLKAKALDTKQNKVQLR 180
DB 121 IVEYEGDVAXXXXXXIEHVESVPEPLPXYYIDPOLFTKGASXLKAKALDTKQNKVQLR 180
QY 181 GIEKIAQXXXNDVXYITAKPEKVNNDVARGIYKADVQSSYGLVGQGOIVAVADTGLD 240
DB 181 GIEKIAQXXXNDVXYITAKPEKVNNDVARGIYKADVQSSYGLVGQGOIVAVADTGLD 240
QY 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNKXTKGNAPQANLYF 300
DB 241 TGRNDSMHEAFRGKITLALYALGRTNANDTNGHGHVAGSVLGNKXTKGNAPQANLYF 300
QY 301 OSIMDSXGGLGGLPSNLQTLFQOASAGARITNSKGAALVNGAYTTDSRNVDDYRKNDY 360
DB 301 OSIMDSXGGLGGLPSNLQTLFQOASAGARITNSKGAALVNGAYTTDSRNVDDYRKNDY 360
QY 361 TILFAAGNEXPNGGTISAPGTAKMAITVGATENTLRPSFGSYADNINHVAFSSRGPTKDG 420
DB 361 TILFAAGNEXPNGGTISAPGTAKMAITVGATENTLRPSFGSYADNINHVAFSSRGPTKDG 420
QY 421 RIKPDVMAFGTYILSARSLAPDSSFWANHDSKYAVMGTSNATPIVAGNVQLREHFVK 480
DB 421 RIKPDVMAFGTYILSARSLAPDSSFWANHDSKYAVMGTSNATPIVAGNVQLREHFVK 480
QY 481 NRGITPKESLLAALLAGAADYGLGYPNGNCGMGRVTLDKSLNAVYNSSXLSLTSQKAT 540
DB 481 NRGITPKESLLAALLAGAADYGLGYPNGNCGMGRVTLDKSLNAVYNSSXLSLTSQKAT 540
QY 541 YKFTTAGKPLKISLWSDAPASTTASVTLVNDLVLVITAPNGTYVNGVDFXKPKXXND 600
DB 541 YKFTTAGKPLKISLWSDAPASTTASVTLVNDLVLVITAPNGTYVNGVDFXKPKXXND 600
QY 601 GRNNVENVPINXPOSGTYTIEVOAYNVVPVGPQXPSLAIYN 640
DB 601 GRNNVENVPINXPOSGTYTIEVOAYNVVPVGPQXPSLAIYN 640

RESULT 9
US-09-509-814A-1
Sequence 1, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURES:
NAME/KEY: misc feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature

LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 98.3%; Score 3007; DB 2; Length 639;

Beet Local Similarity 100.0%; Pred. No. 2.6e-256; Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 KKKVFLSVLSAAALISTVALNPNPSGAXRXPDLDFKGIQTITDDXXGFSKXQGTGAAPLV 63
3 KKKVFLSVLSAAALISTVALNPNPSGAXRXPDLDFKGIQTITDDXXGFSKXQGTGAAPLV 62
64 ESENVKLLKGLKCLLETTPANNKCHIXQFNPIIETQXLEKXGAKLIDYIPYAYIVE 123
63 ESENVKLLKGLKCLLETTPANNKCHIXQFNPIIETQXLEKXGAKLIDYIPYAYIVE 122
124 YEGDVYXXXXIIEHVSEVPLPYXXIDPOLFTGASGLVRAXXLDLTKQXNKEVQLRGIR 183
123 YEGDVYXXXXIIEHVSEVPLPYXXIDPOLFTGASGLVRAXXLDLTKQXNKEVQLRGIR 182
184 XIAQXXSNDVXYITAKPEYRVMDVARGIVKADVAQSSYGLYGGGQIVAAVADTGLDGR 243
183 XIAQXXSNDVXYITAKPEYRVMDVARGIVKADVAQSSYGLYGGGQIVAAVADTGLDGR 242
244 NDSMHEAFRQKITLALVAGTNNANDTNGGTHVAGSVLNGXTNKGMAQANLVFQSI 303
243 NDSMHEAFRQKITLALVAGTNNANDTNGGTHVAGSVLNGXTNKGMAQANLVFQSI 302

Qy 304 MDSXGIGLPSNLTQTLFSDQXASGARIHTNSWGAIVNGAYTTDSRNVDDYVKNMTLL 363
 Db 303 MDSXGIGLPSNLTQTLFSDQXASGARIHTNSWGAIVNGAYTTDSRNVDDYVKNMTLL 362
 Qy 364 PAAGNEXPNGGTTISAPGTAKNAITVGATENTLRPSFSGYADNINHWAFSSRGPTKGRIR 423
 Db 363 PAAGNEXPNGGTTISAPGTAKNAITVGATENTLRPSFSGYADNINHWAFSSRGPTKGRIR 422
 Qy 424 PDVMAAGTITLSARSSLPADSSFWANHDKRYAMGTSNATPIVAQVNAQLRHFVKNRG 483
 Db 423 PDVMAAGTITLSARSSLPADSSFWANHDKRYAMGTSNATPIVAQVNAQLRHFVKNRG 482
 Qy 484 ITPKPSILKALILAGAADXGLGYPNGNGRVTLLDKSLNVAIVNNESSXILSTSOKATYXF 543
 Db 483 ITPKPSILKALILAGAADXGLGYPNGNGRVTLLDKSLNVAIVNNESSXILSTSOKATYXF 542
 Qy 544 TATAGKPLKISLWMSDAPASTTASVTLVNDLDTITAPNGTXYVGNDFXXPXXNMWGRN 603
 Db 543 TATAGKPLKISLWMSDAPASTTASVTLVNDLDTITAPNGTXYVGNDFXXPXXNMWGRN 602
 Qy 604 NVENPINKPQSGTITIEVQAVNPVGPQXPSLIATVN 640
 Db 603 NVENPINKPQSGTITIEVQAVNPVGPQXPSLIATVN 639

RESULT 10
 US-09-920-954-1
 Sequence 1, Application US/09920954
 Patent No. 6759228
 GENERAL INFORMATION:
 APPLICANT: TAKAIWA, MIKIO
 APPLICANT: OKUDA, MITSUYOSHI
 APPLICANT: SAEKI, KATSUHIISA
 APPLICANT: KUBOTA, HIROMI
 APPLICANT: HITOMI, JUN
 APPLICANT: KAGEYAMA, YASUSHI
 APPLICANT: SHIRATA, SHIGESHI
 APPLICANT: NOMURA, MASAFUMI
 TITLE OF INVENTION: ALKALINE PROTEASE
 FILE REFERENCE: 0327-0832-0PCT
 CURRENT APPLICATION NUMBER: US/09/920,954
 CURRENT FILING DATE: 2001-08-03
 PRIOR APPLICATION NUMBER: 09/509,814
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: PCT/JP98/04528
 PRIOR FILING DATE: 1998-10-07
 PRIOR APPLICATION NUMBER: JP 9-274570
 PRIOR FILING DATE: 1997-06-08
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 1
 LENGTH: 639
 TYPE: PRT
 ORGANISM: Bacillus sp.
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (23)..(23)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (29)..(29)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (32)..(32)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (46)..(46)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (47)..(47)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (53)..(53)
 OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc_feature
 LOCATION: (70)..(70)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (74)..(74)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (89)..(89)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (102)..(102)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (105)..(105)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (128)..(128)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (130)..(130)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (131)..(131)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (132)..(132)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (146)..(146)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (148)..(148)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (160)..(160)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (165)..(165)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (172)..(172)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (183)..(183)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (187)..(187)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (188)..(188)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (189)..(189)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (194)..(194)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (286)..(286)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (306)..(306)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (324)..(324)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature
 LOCATION: (369)..(369)
 OTHER INFORMATION: Xaa is any amino acid
 NAME/KEY: misc_feature

LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-920-954-1

Query Match 98.3%; Score 3007; DB 2; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.6e-256;

Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 KKKVPLSTLSAAAIISTVALXNPSSAGXARXFDLPFGKIQTTTDXKGFSGKXGTGAALV 63
3 KKKVPLSTLSAAAIISTVALXNPSSAGXARXFDLPFGKIQTTTDXKGFSGKXGTGAALV 62
64 ESENVKXKXGKXKXKLETPANNKLIHQFNGPILEETKXLEXTGAKILDIYIPDYAIVE 123
63 ESENVKXKXGKXKXKLETPANNKLIHQFNGPILEETKXLEXTGAKILDIYIPDYAIVE 122
124 YEGDVYKXXXXIENHVESVEPYLPXYXIDPOLFTKGASXLYVKXALDTRKXNKEVQLRGIE 183
123 YEGDVYKXXXXIENHVESVEPYLPXYXIDPOLFTKGASXLYVKXALDTRKXNKEVQLRGIE 182
184 XIAQXXXXNDVYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQGLVAVADTGLDTR 243
183 XIAQXXXXNDVYITAKPEYKVMNDVARGIVADVAQSSYGLYGGQGLVAVADTGLDTR 242
244 NDSNHEAPRGKITLALYALGRTNNANDTNGHGTTHVAGSYLGXGTNKGAPANIVFOSI 303
243 NDSNHEAPRGKITLALYALGRTNNANDTNGHGTTHVAGSYLGXGTNKGAPANIVFOSI 302
304 MDSXGGLGLPSNLTOTLPSQASAGARITHNSGMAVNGAYTTDSRNVDDYRKNDMTL 363
303 MDSXGGLGLPSNLTOTLPSQASAGARITHNSGMAVNGAYTTDSRNVDDYRKNDMTL 362
364 FAAGNEXNGGTISAPGTAKNAITVAGATENLAPSGSYADNINHYAQSRSRPTDGRX 423
363 FAAGNEXNGGTISAPGTAKNAITVAGATENLAPSGSYADNINHYAQSRSRPTDGRX 422
424 PDVMAPEGXIIISARSLAPDSSFMANHDSKVIYNGGTSNATPIYVAGNVAAQLREHYVKRG 483
423 PDVMAPEGXIIISARSLAPDSSFMANHDSKVIYNGGTSNATPIYVAGNVAAQLREHYVKRG 482

484 ITPKPSLLKALINGADGGLGYPNGNOGMRVTLDKSLNVAVYNESSXLSQKATYX 543
483 ITPKPSLLKALINGADGGLGYPNGNOGMRVTLDKSLNVAVYNESSXLSQKATYX 542
544 TATAGKPLKISLVMSDAPASTTASVTLVNDLVTITAPNGTXVYVNDPXXKXXNWDGRN 603
543 TATAGKPLKISLVMSDAPASTTASVTLVNDLVTITAPNGTXVYVNDPXXKXXNWDGRN 602
604 NVENVFINXPOSGTYTTEVQAYNPVGPQKPSLAIVN 640
603 NVENVFINXPOSGTYTTEVQAYNPVGPQKPSLAIVN 639

RESULT 11

US-08-873-479-42
Sequence 42, Application US/08873479
Patent No. 5891701
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESS: No. 5891701 of No. 5891701th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873, 479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Agilis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 641 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-873-479-42

Query Match 89.0%; Score 2722.5; DB 1; Length 641;
Best Local Similarity 82.9%; Pred. No. 3e-231;

Matches 532; Conservative 40; Mismatches 67; Indels 3; Gaps 2;

1 MRXK--KKVPLSVSAAAIISTVALXNPSSAGXARXFDLPFGKIQTTTDXKGFSGKXGTGA 58
1 MRXKSKRVPLSVLSVALLSSVALLSSSTIGANNFELDFGIERLTLEKATATQKGTGK 60
59 AAFVSEENVKXKXKXKLETPANNKLIHQFNGPILEETKXLEXTGAKILDIYIPDY 118
61 ASFLVNSENVKIPKSIQKGLVVPADNKLIVQFDPGPILEETQQLLEFTGAKILDIYIPDY 120
119 AYIYVEEDVYKXXXXIENHVESVEPYLPXYXIDPOLFTKGASXLYVKXALDTRKXNKEVQ 178
121 AYIYVEDVYKVAVTNALHLESVEPYLPXYXIDPOLFTKGASXLYVKXALDTRKXNKEVQ 180


```

QY 179 LGIEKIXIAOXXXXNDXYITTAKEKYKAMDVAIGIKADVAOSXYGLYGGGQIVAAADG 238
Db 181 LRGLEQIAOYATNNDDVLYTTPKREYELNDVARGIKADVAQNPFGLYGGQIVAAADG 240
QY 239 LDYGRNDSMHBAFRGKITALYALGATNNANDTNGHGTIVAGSVLNGKXNNKMAPOANL 238
Db 241 LDYGRNDSMHBAFRGKITALYALGATNNANDPNGHGTIVAGSVLNG - ATNNKMAPOANL 239
QY 299 VPOSIMDSYGGIGGUPSNITQTLPSQKXSNAGARIHTNMSKGAUVNAGATTTDSRNDVYRK 358
Db 300 VPOSIMDSQGGIGGUPANIQTLPSQKYSAGARIHTNMSKGAUVNAGATTTDSRNDVYRK 359
QY 359 DWTILPFAAGNEXPNGGTTISAPGTAKAKIIVGATENLRPFGSYADYNINHVAQSSRGPTK 418
Db 360 DWTILPFAAGNEBPGSTISAPGTAKAKIIVGATENLRPFGSYADYNINHVAQSSRGPTK 419
QY 419 DGRIRPDVMAPGTXIISANSIAPDSSFPANHDSKAYNMGTSMATPIVAGNVAQIREHF 478
Db 420 DGRIRPDVMAPGTYIISARSSILAPDSSFPANHDSKAYNMGTSMATPIVAGNVAQIREHF 479
QY 479 VKNRGITTPRSLKKAALIGAADXGIGYNGNGQWGRVTLDTSLNAYNVAWSESXLSJSOK 538
Db 480 VKNRGITTPRSLKKAALIGAADVGLGFPNGNGQWGRVTLDDSLNVAFWNETISPLSTSOK 539
QY 539 AITXETATAGKPLKISLVNSDAPASTTASVTLVNDJDLVITAPNGTXYYGNDPXXEXKXN 598
Db 540 AITXETAGKPLKISLVNSDAPGSTTASJTLVNDJDLVITAPNGTKYYGNDPTAYDNN 599
QY 599 WDGRNNVENVFVNXPOSCTYTIIEVQAYNVPVGPQXSLAIYN 640
Db 600 WDGRNNVENVFVINAPOSCTYTIIEVQAYNVPVSPQTSILAIYH 641

```

```

1      RESULT 12
2      US-08-873-479-43
3      ; Sequence 43: Application US/08873479
4      ; Patent No. 5891701
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Sloma, Alan
7      ; APPLICANT: Lynne, Christiamson
8      ; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
9      ; TITLE OF INVENTION: Having Protease Activity
10     ; NUMBER OF SEQUENCES: 57
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: No. 589170101 No. 5891701disk of No. 5891701th America
13     ; STREET: 405 Lexington Avenue
14     ; CITY: New York
15     ; STATE: NY
16     ; COUNTRY: USA
17     ; ZIP: 10174
18     ;
19     ; COMPUTER READABLE FORM:
20     ; MEDIUM TYPE: Diskette
21     ; COMPUTER: IBM Compatible
22     ; OPERATING SYSTEM: DOS
23     ; SOFTWARE: FastSeq for Windows Version 2.0
24     ; CURRENT APPLICATION DATA:
25     ; APPLICATION NUMBER: US/08/873,479
26     ; FILING DATE: 12-JUN-1997
27     ; CLASSIFICATION: 530
28     ; ATTORNEY/AGENT INFORMATION:
29     ; NAME: Agtis, Cheryl H
30     ; REGISTRATION NUMBER: 34,086
31     ; REFERENCE/DOCKET NUMBER: 5251.000-US
32     ; TELECOMMUNICATION INFORMATION:
33     ; TELEPHONE: 212-867-0123
34     ; TELEFAX: 212-878-9655
35     ;
36     ; TELEX:
37     ; INFORMATION FOR SEQ ID NO: 43:
38     ; SEQUENCE CHARACTERISTICS:
39     ; LENGTH: 635 amino acids
40     ; TYPE: amino acid
41     ; STRANDEDNESS: single
42     ; TOPOLOGY: linear
43     ;

```

US-08-873-479-43

Query Match	79.2%;	Score 2423.5;	DB 1;	length 635;
Best Local Similarity	73.1%;	Pred. No. 6.5e-205;		
Matches 468;	Conservative 68;	Mismatches 99;	Indels 5;	Gaps 4;

Qy	1	MRXKKVVLSTVLSAAALISTVALKNKSNPSGXAKXFLDIDFQIGITOTTDXXGFSKXOXRGA	60
Db	1	MKGKKRVLSTVVASAAILIASVWSSPTGSA--DFQVNFVGK-SLENABLVPKISSGEAS	57
Qy	61	FLVSENVKLAKKGLKPLKLPVPAANKLHIQFNGPILBEETKQXLEXTGAKLIDYIDYAY	120
Db	58	FLVPTENINI PKGIQKLEAVQKQNELIYPTGTPISEBERKGLSEBLGVSILDYVDYAF	117
Qy	121	IVYEGDVAXXXXXLHYESVEPYLPYXXIPDQLTTKASXLVKAXALDTKKQNEVQLR	180
Db	118	IVQVSG-ATKNISTHASEVNWQFLPLKYPIDELLTKGASQLVQAAILNTKHNKMKKT	176
Qy	181	GIEKIAQXXXXNDVYITAKEPEYKWNVDVARIYADVAQSSYGLYGQCIYAAVADTGLD	240
Db	177	GLDELVOYAAANNDDVYITSPKEPEYELMNDVARGIYADVAQNNYGLYGQCIYAAVADTGLD	236
Qy	241	TGRNDSMHEAPRGKITLVALGRTNNAANDTNGHGTTHVAGSVLGNKTKNGMAPQANLVP	300
Db	237	TGRNDSMHEAPRGKITLVALGRTNNAASDPNGHGTTHVAGSVLGN-ALNKGMAPQANLVP	295
Qy	301	QSIDMSXGGLGGLPSNLTQTLPSQAXSAGARITHNSMGAAVNAAYTTDSRNVDDYAKNM	360
Db	296	QSIDMSXGGLGGLPSNLTPLFSQAMNAGARITHNSMGAPVNCAYTTANSQVDEYANNM	355
Qy	361	TILPAGNEXPNGGTISAPGTAKAALTVGATENLPSFGSYADNINHVQAFSSRGPTKQD	420
Db	356	TVLPAAGNGBNSGTTISAPGTAKAALTVGATENYRPSFGSIDNPNHIAQFSSRGATDQ	415
Qy	421	RIRKDWAAPGTXILISARSLSLAPDSSPFMANHDSKYAMGTSMATPIVAGNAVLRRHPK	480
Db	416	RIRKDVYAPGTFILISARSLAPDSSPFMANYSKYAMGTSMAATPIVAGNAVLRRHFAT	475
Qy	481	NRGITPXRSLIKAALIAAGAXGLGYPANGNQMGVTLDKSLNVAVVNSSXSLSTSQKAT	540
Db	476	NRGITPXRSLIKAALIAAGATDVGGLGYPBGDQMGVTLDKSLNVAVVNEATLATAQCKAT	535
Qy	541	YKFTATAGKPLIKISLWSADAPASTTASVTLVNDLVLITAPNGTXYYVGNDFYXPKXXND	600
Db	536	YSFOAQAGKPLIKISLWTDADAGSTTASVTLVNDLVLITAPNGKRVGNDFSYPYDNNM	595
Qy	601	GRNNVENYFNKXPOSGTTIEVOANVNVGPPXPSLAIYN	640
Db	596	GRNNVENYFNAPQSGTYLIEVQANVNSGPPRFSLAIYN	635

RESULT 3
US-09-985-689A-1
Sequence 1, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483USO
CURRENT APPLICATION NUMBER: US/09/985,689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.7e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVARGIVKADVAAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITALVALGRTN 266
DB 1 NDVARGIVKADVAAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITALVALGRTN 60
QY 267 NANTNGHGTAVAGSVLGNKGTNKGMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSSQAXS 326
DB 61 NANTNGHGTAVAGSVLGNKGTNKGMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSSQAXS 120
QY 327 AGARHTNSWGAANVGAATTTDSRVNDVYRKNDMTILPAAGNEXPNGGTTISAPGTAKXAI 386
DB 121 AGARHTNSWGAANVGAATTTDSRVNDVYRKNDMTILPAAGNEXPNGGTTISAPGTAKXAI 180
QY 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTXILSARSSILAPDSSF 446
DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTXILSARSSILAPDSSF 240
QY 447 WANHSKAYWGTSMATPIVAGNVAQLREHFVKNRGITTPKPSILKALAGAADVGLGY 506
DB 241 WANHSKAYWGTSMATPIVAGNVAQLREHFVKNRGITTPKPSILKALAGAADVGLGY 300
QY 507 PNGNGMGRVTLDSKLNVAAYNBSXLSQKATYXFTATGKPLKISLWSDAPASTTA 566
DB 301 PNGNGMGRVTLDSKLNVAAYNBSXLSQKATYXFTATGKPLKISLWSDAPASTTA 360
QY 567 SVTLVNDLDVITAPNGTYVGNDFXXPKXKXNDGRNVENVFINXPOSGTYTIEVOAYN 626
DB 361 SVTLVNDLDVITAPNGTYVGNDFXXPKXKXNDGRNVENVFINXPOSGTYTIEVOAYN 420
QY 627 VPVGPOXFSIAIYN 640
DB 421 VPVGPOXFSIAIYN 434

RESULT 14
US-09-985-689A-2
Sequence 2, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:

APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGETAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 70.4%; Score 2155; DB 2; Length 434;
Best Local Similarity 96.3%; Pred. No. 1.7e-181;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 207 NDVARGIVKADVAAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITALVALGRTN 266
DB 1 NDVARGIVKADVAAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITALVALGRTN 60
QY 267 NANTNGHGTAVAGSVLGNKGTNKGMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSSQAXS 326
DB 61 NANTNGHGTAVAGSVLGNKGTNKGMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSSQAXS 120
QY 327 AGARHTNSWGAANVGAATTTDSRVNDVYRKNDMTILPAAGNEXPNGGTTISAPGTAKXAI 386
DB 121 AGARHTNSWGAANVGAATTTDSRVNDVYRKNDMTILPAAGNEXPNGGTTISAPGTAKXAI 180
QY 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTXILSARSSILAPDSSF 446
DB 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTXILSARSSILAPDSSF 240
QY 447 WANHSKAYWGTSMATPIVAGNVAQLREHFVKNRGITTPKPSILKALAGAADVGLGY 506
DB 241 WANHSKAYWGTSMATPIVAGNVAQLREHFVKNRGITTPKPSILKALAGAADVGLGY 300
QY 507 PNGNGMGRVTLDSKLNVAAYNBSXLSQKATYXFTATGKPLKISLWSDAPASTTA 566
DB 301 PNGNGMGRVTLDSKLNVAAYNBSXLSQKATYXFTATGKPLKISLWSDAPASTTA 360
QY 567 SVTLVNDLDVITAPNGTYVGNDFXXPKXKXNDGRNVENVFINXPOSGTYTIEVOAYN 626
DB 361 SVTLVNDLDVITAPNGTYVGNDFXXPKXKXNDGRNVENVFINXPOSGTYTIEVOAYN 420
QY 627 VPVGPOXFSIAIYN 640
DB 421 VPVGPOXFSIAIYN 434

RESULT 15
US-09-985-689A-6
Sequence 6, Application US/09985689A
Patent No. 6803222
GENERAL INFORMATION:

APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGETAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985,689A
PRIOR FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin version 3.1
SEQ ID NO 6
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 68.1%; Score 2082; DB 2; Length 434;
Best Local Similarity 91.5%; Pred. No. 4.6e-175;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 207 NDVARGIVKADVAAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITALVALGRTN 266
DB 1 NDVARGIVKADVAAQSSYGLYGQGOIVAVADTGLDGTGRNDSMHBAFRGKITALVALGRTN 60

```

Qy 267 NANDTNGHGTVAAGSVLGNKXTNKGMARQANLVFQSIIMDSXGIGGLPSNLQTLPSQAYS 326
Db 61 NANDPNGHGTVAAGSVLGNKXTNKGMARQANLVFQSIIMDSXGIGGLPSNVSTLPSQAYS 120
Qy 327 AGARIHNTSMGAIVNGAYTTDSRNVDYRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
Db 121 AGARIHNTSMGAPVNGAYTTDSRNVDYRKNDMAVLFAAGNEXPNGGTISAPGTAKNAI 180
Qy 387 TVGATENTLRFPSGSIADNINHVAQFSSRGPTDGRIKPDVMAPGTIXILSARSSLAPDSSF 446
Db 181 TVGATENTLRFPSGSIADNINHVAQFSSRGPTDGRIKPDVMAPGTIFLSARSSLAPDSSF 240
Qy 447 WANHDSKXAYMGSTMAPIVAGNVAQLREHFPVKRGTIPKPSLLKALILAGADKGLGY 506
Db 241 WANHDSKXAYMGSTMAPIVAGNVAQLREHFPVKRGTIPKPSLLKALILAGATDGLGY 300
Qy 507 PNGNOGMRVTLDKSLNVAYNVSSXYLSTOKATYXFTATAGKPLKISLWSDAPASTYA 566
Db 301 PNGNOGMRVTLDKSLNVAFVNETSLSLSTNOKATYSFTAGSGKPLKISLWSDAPASTYA 360
Qy 567 SYTLVNDLDELVTAPNGTXYVGNDFKXPKXXNMDCNNVENYFINXPOS GTTIEVQAYN 626
Db 361 SYTLVNDLDELVTAPNGTXYVGNDFTA PYDNMWDGNNVENYFINAPQSGT YTVVEQAYN 420
Qy 627 VPVGPOXPSLAIYN 640
Db 421 VPQGPAPPSLAIYN 434

```

Search completed: April 7, 2006, 17:44:49
 Job time : 35.0266 secs

Y:

This Page Blank (uspto)

1	276.5	9.0	802	6	US-10-510-386-2	Sequence 2, April
2	274.5	9.0	874	6	US-10-510-386-28	Sequence 28, April
3	274.5	9.0	1047	6	US-10-510-386-200	Sequence 200, April
4	242.5	7.9	379	7	US-11-156-062-23	Sequence 23, April
5	242	7.9	272	7	US-11-020-602-236	Sequence 26, April
6	239	7.8	269	7	US-11-020-602-6	Sequence 6, April
7	229	7.5	362	7	US-11-020-602-2	Sequence 2, April
8	227.5	7.4	274	7	US-11-156-062-14	Sequence 14, April
9	227.5	7.4	275	7	US-11-065-943-54	Sequence 54, April
10	227.5	7.4	275	7	US-11-020-602-3	Sequence 3, April
11	226.5	7.4	274	7	US-11-156-062-12	Sequence 12, April
12	225.5	7.4	274	7	US-11-156-062-5	Sequence 5, April
13	225.5	7.3	274	7	US-11-156-062-4	Sequence 4, April
14	221.5	7.2	274	7	US-11-156-062-10	Sequence 10, April
15	220.5	7.2	274	7	US-11-156-062-6	Sequence 6, April
16	220.5	7.2	274	7	US-11-156-062-8	Sequence 8, April
17	220.5	7.2	275	7	US-11-020-602-4	Sequence 4, April
18	217.5	7.1	274	7	US-11-156-062-16	Sequence 16, April
19	217.5	7.1	274	7	US-11-156-062-18	Sequence 18, April
20	216.5	7.1	274	7	US-11-156-062-2	Sequence 2, April
21	210.5	6.9	1052	7	US-11-020-602-208	Sequence 208, April
22	205.5	6.7	1432	6	US-10-510-386-228	Sequence 218, April
23	198.5	6.5	280	7	US-11-020-602-209	Sequence 209, April
24	197.5	6.5	1647	7	US-11-052-554A-260	Sequence 260, April
25	179	5.9	740	7	US-11-096-568A-24714	Sequence 24714, April

26	1.79	5.9	777	7	US-11-096-568A-24713	Sequence 24713, A
27	1.79	5.9	790	7	US-11-096-568A-24712	Sequence 24712, A
28	166.5	5.4	591	6	US-10-510-386-22	Sequence 22, Appl
29	160	5.2	722	7	US-11-096-568A-11863	Sequence 11863, A
30	160	5.2	757	7	US-11-096-568A-11862	Sequence 11862, A
31	160	5.2	798	7	US-11-096-568A-11861	Sequence 11861, A
32	152	5.0	733	7	US-11-096-568A-24028	Sequence 24028, A
33	152	5.0	753	7	US-11-096-568A-24027	Sequence 24027, A
34	152	5.0	764	7	US-11-096-568A-24026	Sequence 24026, A
35	151	4.9	659	7	US-11-096-568A-17896	Sequence 17896, A
36	151	4.9	671	7	US-11-096-568A-17895	Sequence 17895, A
37	151	4.9	791	7	US-11-096-568A-17894	Sequence 17894, A
38	150	4.9	791	7	US-11-096-568A-31251	Sequence 31251, A
39	146.5	4.9	794	7	US-11-218-966-2	Sequence 2, Appl
40	148.5	4.9	820	6	US-10-821-234-1176	Sequence 1176, Appl
41	143	4.7	672	7	US-11-096-568A-23983	Sequence 23983, A
42	143	4.7	680	7	US-11-096-568A-23982	Sequence 23982, A
43	134	4.4	617	6	US-10-872-528-3	Sequence 3, Appl
44	132.5	4.3	617	6	US-11-096-568A-31252	Sequence 31252, A
45	130	4.2	418	7	US-11-096-568A-23984	Sequence 23984, A

ALIGNMENTS

```

RESULT 1
US-10-510-386-2
; Sequence 2, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 802
; TYPE: PRY
; ORGANISM: Bacillus licheniformis
US-10-510-386-2

```

Query Match 9.0%; Score 276.5; DB 6; Length 802;

Matches 160; Conservative 82; Mismatches 226; Indels 265; Gaps 31;

Matches 160; Conservative 82; Mismatches 226; Indels 265; Gaps 31;

127 DVXSXXXI EHVESVEPYLPXYXI DPQLFTKGASXLVKALDT---KQXNKEVQLRGIE 183

49 DVTSDKQTTIVIVELKEKSLABAKADGEKÖTKASLKTARSKALTKLICKAKVNBREYD-RVFS 107

184 XIAQXXXSNDVXYITAKPEYKM-----NDVARGIVKAD 217

108 GPSMKLPASEIPKLLAVKEVKAVYPNATYKPDSEVKGDVTLAADAIPPQMDKSAPFIGAD 167

218 VAQSSYGLYGQGIIVADTGLDTRND-SSMHEAFRGK--ITALYALGRTNNDNTNG- 273

168 QAWKS-GYTGKGIKVAIDTGVDTYHPDLKNNFGPYKGYDFVNDNDYDPQETPTGCDPRGGA 226

274 --HGTHVAGSVLGNXTNKGMAPQANLVQSIMDSXGGGLPSNLQTLPSQAXSAGARI 331

227 TDHGTHVAGTIAANGQI-KGVAPBEATLLAYRVLGPGG--SCTTENVIAGIBKAVADGAKV 283

332 HTNSWGA VNGA - YTTDSRNVDYVRKNDMTILPAGNEXPNGGTISAPGTAKNAITVGA 390

284 MNLSLGNSLNSPDYATSI--ALDMAMAEGVAVTISNGNSGPENNTVGSPTSRVAISVGA 341

391 TE-----NLKPSFGS-----400

Db 342 SOLPYNEYSVTLPSYSKAWMGYOEKDELEALNGEVELVEAGLQADDPSGKDVKGVA 401
Qy 401 -----YADININ----- 407
Db 402 VIORGVIPEVDKAKENAKAGAI GAVIYNNATGEIEANMGMAVPTVKLSKEEGELVQOI 461
Qy 408 -----VAOFSSRGPTKD-GRIKPDVMAPTGYILSARSLAPDSF 446
Db 462 KEGKHSVVPFLDKKGLBETTLASFSSRGVMDTNMIKPDVSAFGVNIYSTIPT----- 514
Qy 447 WANHDSK---YAWMGTSMAPTIYAGNVAQLREHFVKNRGITP--KPSLKAALLIAGAA 500
Db 515 ---HDPKRPYVGSGKQGTSMASPHVAAGTAALIKQ-----AKPDMTEQIKGVLMNTAE 564
Qy 501 ---DXGLGYPMGNGMGKRVTLDKSLNV-----AVYNESS 531
Db 565 KLTBNGKPLPNTQAGSIRIMEALKASSIVTPGSHSYGTFKDKKQKOTKQKATPISN- 623
Qy 532 XLSTSQKA---TYXFTATAGKPLKISLWSD--APASTT--ASVTLVNDLIVITAPN 582
Db 624 -LSSHRKAYOLEYSFKGTG---ITVSGTERVVVPANOTGAAAKVTYNSAKTKAGTYE 677
Qy 583 GTXYGVNDPKXXPKXXNMDSGRNVE--NVFINXP-----QSGTTTIEVQAY 625
Db 678 GTVYIIR-----DGRKVAERIPLLIYKEPDYPRVTSVTEBPAGAKQAGYITE--AY 725
Qy 626 NVPVGPOKFXSLAI 638
Db 726 -LPGAEELAPLV 737

RESULT 2

US-10-510-386-28
; Sequence 28, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Raemussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-28

Query Match 9.0%; Score 274.5; DB 6; Length 874;
Best Local Similarity 23.7%; Pred. No. 1.4e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;
Qy 14 AAATLSTVALKNPAGAKRFDLPKGIQTDDXXGFSKQXQTGAAP-----LV 63
Db 213 AAPFISAKMLFNKSMRDSVYKDL---VOLKPLPADVAKNGQVSAYSTINDDKAISRLY 268
Qy 64 ESENVKLLKGLXKKLETVPANNKLHIXQFNPGPILEETKQKLEXTG-AKILDIYIPDYAYIV 122
Db 269 ETAAASVPEPLKQDDQVAKD--IGIEQLTG---SKVSAVLEKGMATASSAPERNRIV 323
Qy 123 EYE-----GDVYKXXXXIEHVESVPE-----YLPYXXID-PQLFTGASXLVAKAALD 169
Db 324 KLKGGKPKGSPKSKAQG-SGVQALEPLGSKTIAFKDMYVEMKESRSRGFRAAAKQYQAA 382
Qy 170 TKQKNKEVQALGIEIXIAG-XXXSNDVXYITAKPEYKMNVDVARGIVKADVAQSSYGLYQG 228
Db 383 ASKIAKMEVEFVEQVOYEALSRDTQY---PYQWLSKNNGKNRAANDIOFEOLOKLMK 439

Qy 229 GQ-----IVAVADTGLDTRGNDSSMHEAFRGKI---TALVALGRTNNAINDTNGHGTIVAG 280
Db 440 GKXKQVYIAVVDVGVDHTLADLS-----GSVKKDEBYNVVGTATADMDNGHGTIVAG 493
Qy 281 SVLG---NGXTNKGMAPOANI VPOSINDSKGGLGSLBSNIOQLTFSQAXSAGARIHTNSWG 337
Db 494 IIAAQODHFEMAGINNAKILPVKVLDSG--SGDTEQIANGIITYADHGAKYINISLIG 551
Qy 338 AAVNGAYTDBSRANDVDYKRNKDMITLPAAGEXHNGGITSAPGAKNAITGATENLRPS 397
Db 552 ---GPSRVMEYALKTAASKNVITVAATGNDVGS--EISTPASSKTTLSVGAITNNL-- 602
Qy 398 FGSYADININHAQFSSRGPTKDGRIKPDVMAPTGYILSARSLAPDSFMANHDSKYAYM 457
Db 603 -----DLVSDISNNGKL-----DMVAPGTDI-----PSLVDPDN-----VTYM 636
Qy 458 GGTSMAPPTIYAGNVAQLREHFVKNRGITPSPSLKKAALLIAGADXLG--GYPNGN---- 510
Db 637 SGTSMALPHVAAAAGLL-----LSQNPISLAKPKQIASILLETADVAFEEQDNPNPDYDLD 691
Qy 511 -----QGMKRVTLDKSLNVAAYVNESSKLSQKATYXFTATAGKPLKI 553
Db 692 IEPAQIPGDFVSGMKRLNVFHAASVPELMKQVHPVLRHTAVTGAFTAGSGVTYKI 747

RESULT 3

US-10-510-386-200
; Sequence 200, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Raemussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 200
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-200

Query Match 9.0%; Score 274.5; DB 6; Length 1047;
Best Local Similarity 23.7%; Pred. No. 1.8e-13;
Matches 141; Conservative 85; Mismatches 253; Indels 117; Gaps 25;
Qy 14 AAATLSTVALKNPAGAKRFDLPKGIQTDDXXGFSKQXQTGAAP-----LV 63
Db 213 AAPFISAKMLFNKSMRDSVYKDL---VOLKPLPADVAKNGQVSAYSTINDDKAISRLY 268
Qy 64 ESENVKLLKGLXKKLETVPANNKLHIXQFNPGPILEETKQKLEXTG-AKILDIYIPDYAYIV 122
Db 269 ETAAASVPEPLKQDDQVAKD--IGIEQLTG---SKVSAVLEKGMATASSAPERNRIV 323
Qy 123 EYE-----GDVYKXXXXIEHVESVPE-----YLPYXXID-PQLFTGASXLVAKAALD 169
Db 324 KLKGGKPKGSPKSKAQG-SGVQALEPLGSKTIAFKDMYVEMKESRSRGFRAAAKQYQAA 382
Qy 170 TKQKNKEVQALGIEIXIAG-XXXSNDVXYITAKPEYKMNVDVARGIVKADVAQSSYGLYQG 228
Db 383 ASKIAKMEVEFVEQVOYEALSRDTQY---PYQWLSKNNGKNRAANDIOFEOLOKLMK 439
Qy 229 GQ-----IVAVADTGLDTRGNDSSMHEAFRGKI---TALVALGRTNNAINDTNGHGTIVAG 280
Db 440 GKXKQVYIAVVDVGVDHTLADLS-----GSVKKDEBYNVVGTATADMDNGHGTIVAG 493
Qy 281 SVLG---NGXTNKGMAPOANI VPOSINDSKGGLGSLBSNIOQLTFSQAXSAGARIHTNSWG 337

Db 494 IIAAODNHFMSAGINAVAKILPVKYLDSGG--SGDTEQIANGIIYAADHAKVINLSIG 551
 Qy 338 AAVNGAYTTDSNNVDYVAKNMTLLFAAGNEXPNCGTSAPGTAKNALTVCATENTLRSS 397
 Db 552 ---GPYSVMEYALKYAASKVVTIYAAGNDGVS--EISYPASSKYTLISVGAITNML--- 602
 Qy 338 FGSYADININHVAFSGSRGPTKGRIRKPDVMAFGTXILSARSSILAPDSFMANHDSKYAM 457
 Db 603 ---DLVSDYNSYKGL-----DNVAPGTDL---PSLVPDGN-----VTYM 636
 Qy 458 GGTSMATPIVAGNVAAQLREHFPVKNNGITPKPSILAKAIIAGAADKGL---GYPNGN--- 510
 Db 637 SGTSMAAPHVAAAAGL-----LSQNPISLKPQIASILTETTADVAFEEQDNPNPDYDLD 691
 Qy 511 -----QGMGRVTLDSKLSLVAVVNESSKLSYKTXKTAAGRPKX 553
 Db 692 IEPPAQIPGYDFVSGWGRINLVFAASVFEIANKVHPVLRHTAVGTAKSGVTYK 747

RESULT 4

US-11-156-062-23
 ; Sequence 23, Application US/11156062
 ; Publication No. US2005028173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wieland, Susanne
 ; APPLICANT: Polanyi-Bald, Laura
 ; APPLICANT: Prueser, Inken
 ; APPLICANT: Stehr, Regina
 ; APPLICANT: Mauser, Karl-Heinz
 ; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
 ; FILE REFERENCE: HENK-0134 / H5698
 ; CURRENT APPLICATION NUMBER: US/11/156,062
 ; CURRENT FILING DATE: 2005-06-17
 ; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
 ; PRIOR FILING DATE: 2003-12-20
 ; PRIOR APPLICATION NUMBER: DE 102 60 903.9
 ; PRIOR FILING DATE: 2002-12-20
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 23
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: Bacillus licheniformis
 ; US-11-156-062-23

Query Match 7.9%; Score 242.5; DB 7; Length 379;

Best Local Similarity 23.4%; Pred. No. 1.5e-11;
 Matches 116; Conservative 58; Mismatches 166; Indels 155; Gaps 18;

Qy 1 MRKKKVLSTLSAAIISTVALKNPSAKX--ARKFDLDFKGIQTTTXXGFSKQXQGA 58
 Db 1 MRRKKSFWLGMILTALFVTAFAFSDSASAPAKRVEKDY-----IVGFKSGVKTAS 52
 Qy 59 AAFVSESNVGLXKGLKLETVPANNNKLAHIQFNGPILEETKOKLXTGAIIIDYIPDY 118
 Db 53 ---VKQIITIEGSKVDK-----QFR--IINAARKLIDREALKLEVKNDDPV 93
 Qy 119 AYIYEGSDVAXKXKXIEHVESVEPYLPYXKIDPOLFTKGASXLVAXALDTRKXNKEVQ 178
 Db 94 AYVER-----DHVAHALAQVTPYGP----- 114
 Qy 179 LRGEIXIAKXKXNDVXYITAKPEYKVNNDVARGIVKADVAGSYGLYGQGIYAVADTG 238
 Db 115 -----LIKDKVQAQ--GFKGANVYKVAVLDTG 139
 Qy 239 LDTGRDSSMHEAFRGKITAYALGRTNNANDTNGHIVAGSV--LNGXKTNKGMAPQA 296
 Db 140 IQASHPDLNVVGG-----ASFVAGEAYN--TDGNHGHVAGTVALADNTTGVLGVAASV 192
 Qy 297 NLVFSINDXKGLGLPSNLQTLFSQAXSAGARIHTHSWGAANVAGAVTTDSRNVDYR 356
 Db 193 SLVAVKVLNNSG--SGSYGIVSGIEWATNTGMDVINNSLGGA---SGSTAMKQAVNDAY 247

Qy 357 KNDMTILFAAGNEXPNG--GTISAPGTAKNALTVCATENTLRSPFSYADININHVAFSSR 414
 Db 248 ARGVVVAAGAAGSSSGSTNTTIGYPAKIDSVTAAGA-----VDSNRSRFSVV 296
 Qy 415 GPTKGRIRKPDVMAFGTXILSARSSILAPDSFMANHDSKYAMGTSMAPIVAGNVAA-- 472
 Db 297 G-----ALEVMAGAGVYSTYPT-----NTYATLNGISMAFHVAGAAALI 338
 Qy 473 ---QLREHFPYKRN 482
 Db 339 LSKHPLSAGVYRN 353

RESULT 5

US-11-020-602-236
 ; Sequence 236, Application US/11020602
 ; Publication No. US20060024764A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Batell, David
 ; APPLICANT: Harding, Fiona
 ; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
 ; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
 ; FILE REFERENCE: GCS27C2
 ; CURRENT APPLICATION NUMBER: US/11/020,602
 ; CURRENT FILING DATE: 2004-12-22
 ; PRIOR APPLICATION NUMBER: US 09/500,135
 ; PRIOR FILING DATE: 2000-02-08
 ; PRIOR APPLICATION NUMBER: US 09/060,872
 ; PRIOR FILING DATE: 1998-04-15
 ; NUMBER OF SEQ ID NOS: 240
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 236
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURES:
 ; OTHER INFORMATION: Description of Artificial Sequence: Hybrid of
 ; OTHER INFORMATION: Bacillus lichenus and Bacillus amyloliquefaciens
 ; US-11-020-602-236

Query Match 7.9%; Score 242; DB 7; Length 272;

Best Local Similarity 31.9%; Pred. No. 1e-11;
 Matches 84; Conservative 30; Mismatches 101; Indels 48; Gaps 10;

Qy 214 VKADVAGSSYGLYGGQIVAAVDTGLDGRDSSMHEAFRGKITAYALGRTNNANDTNG 273
 Db 11 VQAPRAHNR--GLTSGGVAVAVLDGIST-----HPLNIRGASFPVGBE--STQDNG 61
 Qy 274 HGTHVAGSV--LGNXKTNKGMAPQANLVFOSIMDSXGLGLPSNLQTLFSQAXSAGARI 331
 Db 62 HGTHVAGTIALNNISIGVLGVAFSAELVAVKVLGASG--SGSVSSIAQGLEWAGNNGMHV 119
 Qy 332 HTNSGVAIVNGAYTTDSNNVDYVAKNMTLLFAAGNEXPNG--GTISAPGTAKNALTNG 389
 Db 120 INMSLGSSGSAAL-----PAAYDKAVASGVVVAAGNGTGSSSTVGPGRYPSVIANG 175
 Qy 390 ATENLRPSFGSYADININHVAFSGSRGPTKGRIRKPDVMAFGTXILSARSSILAPDSFMAN 449
 Db 176 A-----VDSNRSRFSVGP-----ELDWAPGVISIGSTLRG----- 208
 Qy 450 HDSKYAMGTSMAPIVAGNVAA 472
 Db 209 --NKYGAYNGTSMASFPVAGAAA 229

RESULT 6

US-11-020-602-6
 ; Sequence 6, Application US/11020602
 ; Publication No. US20060024764A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Batell, David
 ; APPLICANT: Harding, Fiona
 ; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND

```
;; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
;; FILE REFERENCE: GCS27C2
;; CURRENT APPLICATION NUMBER: US/11/020,602
;; CURRENT FILING DATE: 2004-12-22
;; PRIOR APPLICATION NUMBER: US 09/500,135
;; PRIOR FILING DATE: 2000-02-08
;; PRIOR APPLICATION NUMBER: US 09/060,872
;; PRIOR FILING DATE: 1998-04-15
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 269
;; TYPE: PRN
;; ORGANISM: Bacillus lentus
US-11-020-602-6

Query Match      7.8%; Score 239; DB 7; Length 269;
Best Local Similarity 30.7%; Pred. No. 1,8e-11;
Matches 90; Conservative 30; Mismatches 91; Indels 82; Gaps 14;

QY 214 VKADVAQSSYGLYGGQIVAAVADTGLDGRNDSMHEAFRGKITALVALGRTNANDTNG 273
DB 11 VQAPAAHNR-GLTSGGVAVAVLDTGIST-----HPDLNIRGASFPVGP-STDGNG 61
QY 274 HGTAVAGSV--LGNGXTKMGAPQANLVFQSIMDSXGGLGSPSLQTLFQASAGARI 331
DB 62 HGTAVAGTIALNNSTIGVLGVAAPSAELYAVKVLGASG--SGSVSSIAQGLEWAGNGMHV 119
QY 332 HTNSGCA-----AVNGAYTTDSRVNDVVRKNDITLLPAAGNEXPNQSTISAPGTA 382
DB 120 ANLSIGSPSPATLEQANNSA---TSRGV-----LVVAASGNS--GAGSISIPARY 165
QY 383 KNAITVGATE--NLRFSGSTADNINHVAGFSSRGPTDGRKIPDVAPGTIILSARSL 440
DB 166 ANAAVAGATDQNNNNAASFQYAGCL-----DIVAPGVVQSTYRG- 205
QY 441 APDSFMANHDSKAYVMGTSMAPTPIVAGNVA-----OLRHFVYN 481
DB 206 -----STYASLNGTSMATPHVAGAAALVYKQKPNMSNVQIRNH-LKN 246

RESULT 7
US-11-020-602-2  #
; Sequence 2, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Bectell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GCS27C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 382
; TYPE: PRN
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (163)...(163)
; OTHER INFORMATION: Xaa = Asn or Pro
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (164)...(164)
; OTHER INFORMATION: Xaa = Pro or Asn
; FEATURE:
; NAME/KEY: VARIANT
```

```
;; LOCATION: (168)...(168)
;; OTHER INFORMATION: Xaa = Asn or Asp
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (195)...(195)
;; OTHER INFORMATION: Xaa = Ala or Ser
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (196)...(196)
;; OTHER INFORMATION: Xaa = Ser or Ala
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (205)...(205)
;; OTHER INFORMATION: Xaa = Ala or Asp
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (206)...(206)
;; OTHER INFORMATION: Xaa = Asp or Ala
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (265)...(265)
;; OTHER INFORMATION: Xaa = Thr or Ser
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (266)...(266)
;; OTHER INFORMATION: Xaa = Ser or Thr
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (358)...(358)
;; OTHER INFORMATION: Xaa = Gln or Glu
US-11-020-602-2

Query Match      7.5%; Score 229; DB 7; Length 382;
Best Local Similarity 23.6%; Pred. No. 1.7e-10;
Matches 117; Conservative 40; Mismatches 151; Indels 188; Gaps 17;

QY 5 KKVLSTVSAAILITSTVALXNPSAGXARXPLDPKGIQTITDXXGFSKQXQGAAPLVE 64
DB 4 KKVWISLFLPALALFTMAFGSTSAQAG-----KSNGEKRYIVGFQOTMSTGSA----- 54
QY 65 SENVKKXGKLKLETPANNKHLIXQFNPILETQKLEXTAKILDIYIPDAYIYEV 124
DB 55 -----KKQDVISEKGVQKQK--YVDAASATLNEKAVKELKQPSVAYVE- 100
QY 125 EGDVXSKXXIERY-----BSVEPYLPYXTIDPOLFTKGASXLVKAALDTKQAKKEVQL 179
DB 101 -----DHYAHVAQSV-PY-----GVSQ-IKAPAL----- 123
QY 180 RGIEIXAQXXXSNQVXYITAKEPKWMDVARGLVKADVAGSSYGLYGGQIVAAVADTGL 239
DB 124 -----HSGGYTGSNVKVAIVDSGI 142
QY 240 DTGRNDSMHEAFRGKITALVALGRTNANDTNGHGTAVAGSV--LGNGXTKMGAPQAN 297
DB 143 D-----SSHPLDVAGGASVWPSETKXFPQXNHGHTVATVAALNNSIGLVGAPGX 196
QY 298 LVFQSIMDSXGGLGGLSPSLQTLFQASAGARIHTNSMGAAVNGAYTTDSRND----- 352
DB 197 LVAVKVLGXG-----SGQYSWINGIEWALNNNDVINS 232
QY 353 -----DYVRKNDITLLPAAGNEXPNQ--GTISAPGTAKNAITVGATENLRP 396
DB 233 LGPSGSAALKAADVAKVAVGVVVAAGNGXGSSSTVGPGRKPSVIVGA----- 286
QY 397 SFGSYADNINHVAGFSSRGPTKGRKIPDVAPGTIILSARSLAPDSSFMANHDSKAY 456
DB 287 -----VDSNQRASFPSSVGP-----ELDVAPGVISQSTLPG-----NKGCA 323
QY 457 MGTSMATPIVAGNVA 472
DB 324 YNGTSMASPHVAGAA 339
```


RESULT 8
US-11-156-062-14
; Sequence 14, Application US/11156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Sueanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maier, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ IDS NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-14

Query Match 7.4%; Score 227.5; DB 7; Length 274;
Best Local Similarity 30.2%; Pred. No. 1.4e-10;
Matches 85; Conservative 35; Mismatches 108; Indels 53; Gaps 11;
QY 213 IYKAVAGSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 272
DB 10 LKADKVAQ--GPKANVAVAVLDGTGQASHPLDNLVVG-----ASFVAGAYNA-DGN 61
QY 273 GHGTVAGSV--LGNGXTKMGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSGAXSAGAR 330
DB 62 GHGTVAGTVALLDITGVLGVAFPSVSLYAVVNLSSG--SGSYGIVSGIEMATTNGMD 119
QY 331 IHTNSGAAVNGAYTTDSHRVDYVRKNDMTILFAAGNEXPN--GTISAPGTAKNAITV 388
DB 120 VTNMSLGGA---SGSTAMKQAVDNVAVRGVVVAAAAGSSGNTTIGYPAKYDSVLA 176
QY 389 GATNELRPSFGSYADNINHVAFQSSRGPTXGRITPDVMAPTXILSARSSLA PSSFWA 448
DB 177 GA-----VDSNSRASFSSVG-----AELEWAPGAGVSTPT----- 210
QY 449 NHDSKYAVMGTSMAPTIYAGNV-----QLREHFVKNR 482
DB 211 ---NTYATWGTSMASPHVGAALILSKHPNLASQVKNR 248

RESULT 9
US-11-065-943-54
; Sequence 54, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:
; APPLICANT: JESTIN, JEAN-LUC
; APPLICANT: VICHIER-GUERRE, SOPHIE
; APPLICANT: FERRIS, STEPHANE
; TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
; TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES,
; FILE REFERENCE: 266426USOXIP
; CURRENT APPLICATION NUMBER: US/11/065,943
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/787,219
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 54
; LENGTH: 275

; TYPR: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-065-943-54

Query Match 7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;
QY 214 VKADVAQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 273
DB 11 IKAP-ALHSQGYTGNVAVAVDSIGD-----SSHPLKVAAGASWVPSSETNPQDNN 63
QY 274 HGTIVAGSV--LGNGXTKMGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSGAXSAGAR 331
DB 64 HGTIVAGTVALLNNSIGVLGVAFPSASLYAVVNLGADG----- 100
QY 332 HTNSGAAVNGAYTTDSHRVD-----DYVRKNDMTILFAAGNEXPN 372
DB 101 -SGQYSWIIINGIEMALANNMDVYNMSLGPGSSAALKAADVAVASGVVVAAAAGBGS 159
QY 373 G--GTISAPGTAKNAITVGA TENLRPSFGSYADNINHVAFQSSRGPTXGRITPDVMAFG 430
DB 160 GSSSTVGYPGKYPYIAVGA-----VDSNSORASFSSVGP-----ELDVMAFG 202
QY 431 TXILSARSSLA PSSFWANHDSKYAVMGTSMAPTIYAGNV 472
DB 203 VSIQSTLPG-----NKGAYNGTSMASPHVGA 232

RESULT 10
US-11-020-602-3
; Sequence 3, Application US/11020602
; Publication No. US20060024764A1
; GENERAL INFORMATION:
; APPLICANT: Batzell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GC527C2
; CURRENT APPLICATION NUMBER: US/11/020,602
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-11-020-602-3

Query Match 7.4%; Score 227.5; DB 7; Length 275;
Best Local Similarity 28.7%; Pred. No. 1.4e-10;
Matches 81; Conservative 26; Mismatches 92; Indels 83; Gaps 9;
QY 214 VKADVAQSSYGLYGGQIVAVADTGLDGRNDSMHEAFRGKITLALGRTNNANDTN 273
DB 11 IKAP-ALHSQGYTGNVAVAVDSIGD-----SSHPLKVAAGASWVPSSETNPQDNN 63
QY 274 HGTIVAGSV--LGNGXTKMGAPQANLVFQSIMDSXGGLGGLPSNLQTLFSGAXSAGAR 331
DB 64 HGTIVAGTVALLNNSIGVLGVAFPSASLYAVVNLGADG----- 100
QY 332 HTNSGAAVNGAYTTDSHRVD-----DYVRKNDMTILFAAGNEXPN 372
DB 101 -SGQYSWIIINGIEMALANNMDVYNMSLGPGSSAALKAADVAVASGVVVAAAAGBGS 159
QY 373 G--GTISAPGTAKNAITVGA TENLRPSFGSYADNINHVAFQSSRGPTXGRITPDVMAFG 430
DB 160 GSSSTVGYPGKYPYIAVGA-----VDSNSORASFSSVGP-----ELDVMAFG 202

Qy 431 TXILASRLAPDSSFMNHSKYAMGTSMTATPIVAGNVA 472
Db 203 VSIQSTLPG-----NKYAGYNGTSMASPHVAGAAA 232

RESULT 11
US-11-156-062-12

Sequence 12, Application US/11156062
Publication No. US20050281773A1
GENERAL INFORMATION:
APPLICANT: Wieland, Susanne
APPLICANT: Polanyi-Bald, Laura
APPLICANT: Prueser, Inken
APPLICANT: Stehr, Regina
APPLICANT: Maurer, Karl-Heinz
TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
FILE REFERENCE: HENK-0134 / H5698
CURRENT APPLICATION NUMBER: US/11/156,062
CURRENT FILING DATE: 2005-06-17
PRIOR APPLICATION NUMBER: PCT/EP2003/014127
PRIOR FILING DATE: 2003-12-20
PRIOR APPLICATION NUMBER: DE 102 60 903.9
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.3
SEQ ID NO 12
LENGTH: 274
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
US-11-156-062-12

Query Match 7.4%; Score 226.5; DB 7; Length 274;
Best Local Similarity 30.2%; Pred. No. 1.7e-10;
Matches 85; Conservative 34; Mismatches 109; Indels 53; Gaps 11;

Qy 213 IVKADVAQSSYGLYGQGIYAVADTGLDGTGRDSSMHEAFRGKITALYALGRTNNANDTN 272
Db 10 LIRKADKYQAQ-GFKGANVAVVLDTGIOASHPDLNVVG-----ASFVAGEAVN-TDGN 61
Qy 273 GHGTHVAGSV--LNGXTNKGMAPQANLVFQSIINDSGGLGSLPSNLQTLFSSQXSGAR 330
Db 62 GHGTHVAGTVAALNTTGTGLGVAAPSVSLYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119
Qy 331 IHTNSWGAAVGAYTTDSRVNDYVRKNDMTILFPAAGNEXPNG--GTISAPGTAKNAITV 388
Db 120 VINNSLGA---SGSTAKQAVDNAVYARGVYVVAAGNSGSGTNTTIGYPAKYDSVIAY 176
Qy 389 GATENLRPFSGSYADNINNHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSFWA 448
Db 177 GA-----VDSNSNRASFSSVG-----AELEVMAPGAGVSTYPT----- 210
Qy 449 NHDSKYAMGTSMTATPIVAGNVA-----QLREHFYKUR 482
Db 211 ---NTYATNGTSMASPHVAGAAAIIISKPNLSAOVYRNR 248

RESULT 12
US-11-020-602-5
Sequence 5, Application US/11020602
Publication No. US20060024764A1
GENERAL INFORMATION:
APPLICANT: Estell, David
APPLICANT: Harding, Fiona
TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
FILE REFERENCE: GCS27G2
CURRENT APPLICATION NUMBER: US/11/020,602
CURRENT FILING DATE: 2004-12-22
PRIOR APPLICATION NUMBER: US 09/500,135
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: US 09/060,872

;; PRIOR FILING DATE: 1998-04-15
;; NUMBER OF SEQ ID NOS: 240
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 5
;; LENGTH: 274
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-11-020-602-5

Query Match 7.4%; Score 225.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 2.1e-10;
Matches 84; Conservative 36; Mismatches 108; Indels 53; Gaps 11;

Qy 213 IVKADVAQSSYGLYGQGIYAVADTGLDGTGRDSSMHEAFRGKITALYALGRTNNANDTN 272
Db 10 LIRKADKYQAQ-GFKGANVAVVLDTGIOASHPDLNVVG-----ASFVAGEAVN-TDGN 61
Qy 273 GHGTHVAGSV--LNGXTNKGMAPQANLVFQSIINDSGGLGSLPSNLQTLFSSQXSGAR 330
Db 62 GHGTHVAGTVAALNTTGTGLGVAAPSVSLYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119
Qy 331 IHTNSWGAAVGAYTTDSRVNDYVRKNDMTILFPAAGNEXPNG--GTISAPGTAKNAITV 388
Db 120 VINNSLGA---SGSTAKQAVDNAVYARGVYVVAAGNSGSGTNTTIGYPAKYDSVIAY 176
Qy 389 GATENLRPFSGSYADNINNHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSFWA 448
Db 177 GA-----VDSNSNRASFSSVG-----AELEVMAPGAGVSTYPT----- 210
Qy 449 NHDSKYAMGTSMTATPIVAGNVA-----QLREHFYKUR 482
Db 211 ---NTYATNGTSMASPHVAGAAAIIISKPNLSAOVYRNR 248

RESULT 13
US-11-156-062-4

Sequence 4, Application US/11156062
Publication No. US20050281773A1
GENERAL INFORMATION:
APPLICANT: Wieland, Susanne
APPLICANT: Polanyi-Bald, Laura
APPLICANT: Prueser, Inken
APPLICANT: Stehr, Regina
APPLICANT: Maurer, Karl-Heinz
TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PEROXYDOLASE ACTIVITY
FILE REFERENCE: HENK-0134 / H5698
CURRENT APPLICATION NUMBER: US/11/156,062
CURRENT FILING DATE: 2005-06-17
PRIOR APPLICATION NUMBER: PCT/EP2003/014127
PRIOR FILING DATE: 2003-12-20
PRIOR APPLICATION NUMBER: DE 102 60 903.9
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.3
SEQ ID NO 4
LENGTH: 274
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic Construct
US-11-156-062-4

Query Match 7.3%; Score 222.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 3.5e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

Qy 213 IVKADVAQSSYGLYGQGIYAVADTGLDGTGRDSSMHEAFRGKITALYALGRTNNANDTN 272
Db 10 LIRKADKYQAQ-GFKGANVAVVLDTGIOASHPDLNVVG-----ASFVAGEAVN-TDGN 61
Qy 273 GHGTHVAGSV--LNGXTNKGMAPQANLVFQSIINDSGGLGSLPSNLQTLFSSQXSGAR 330
Db 62 GHGTHVAGTVAALNTTGTGLGVAAPSVSLYAVKVLNMSG--SGSYSGIVSGIEMATTNGMD 119

```
QY 331 IHTNSGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 388
DB 120 VINMSLGA--SGSTAKQAVDNAYARGVVVAAAGNSGSGNTNTTGYPAKYDSVLAV 176
QY 389 GATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTXILSARSSLAPDSSFMA 448
DB 177 GA-----VDSNSNRASFSSVG-----AELVMAPOAGVYSTYPT----- 210
QY 449 NHDSKYAYMGTSMATPIYAGNVA-----QLRHFVKNR 482
DB 211 ---NTYATWDTGTSMASPHVAGAAALILSKHPVLSASQVKNR 248

RESULT 14
US-11-156-062-10
; Sequence 10, Application US/11156062
; Publication No. US2005028173A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-10

Query Match 7.2%; Score 221.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 4.3e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

QY 213 IYKADVAOSYGLYQGOIYAVADTGLDTRGNDSSMHEAFRGKITALYALGRTNNANDTN 272
DB 10 LIKADKVAQ--GFKANVYKVALDTGICQASHDNLVVG-----ASFVAGAYNA--DGN 61
QY 273 GHGTHVAGSV--LNGXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFPOAXSAGAR 330
DB 62 GHGTHVAGTVALDNTTGVLAGVAPSVSYAVKVLNSSG--SGSYGIVSGIEMATTNGMD 119
QY 331 IHTNSGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 388
DB 120 VINMSLGA--SGSTAKQAVDNAYARGVVVAAAGNSGSGNTNTTGYPAKYDSVLAV 176
QY 389 GATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTXILSARSSLAPDSSFMA 448
DB 177 GA-----VDSNSNRASFSSVG-----AELVMAPOAGVYSTYPT----- 210
QY 449 NHDSKYAYMGTSMATPIYAGNVA-----QLRHFVKNR 482
DB 211 ---NTYATWDTGTSMASPHVAGAAALILSKHPVLSASQVKNR 248

RESULT 15
US-11-156-062-6
; Sequence 6, Application US/11156062
; Publication No. US2005028173A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
```

```
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H5698
; CURRENT APPLICATION NUMBER: US/11/156,062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-156-062-6

Query Match 7.2%; Score 220.5; DB 7; Length 274;
Best Local Similarity 29.9%; Pred. No. 5.1e-10;
Matches 84; Conservative 35; Mismatches 109; Indels 53; Gaps 11;

QY 213 IYKADVAOSYGLYQGOIYAVADTGLDTRGNDSSMHEAFRGKITALYALGRTNNANDTN 272
DB 10 LIKADKVAQ--GFKANVYKVALDTGICQASHDNLVVG-----ASFVAGAYNA--DGN 61
QY 273 GHGTHVAGSV--LNGXTNKGMAPOANLVFOSIMDSXGGLGGLPSNLQTLFPOAXSAGAR 330
DB 62 GHGTHVAGTVALDNTTGVLAGVAPSVSYAVKVLNSSG--SGSYGIVSGIEMATTNGMD 119
QY 331 IHTNSGAAVNGAYTTDSRNVDDYRKNDMTILFAAGNEXPNG--GTISAPGTAKNAITV 388
DB 120 VINMSLGA--SGSTAKQAVDNAYARGVVVAAAGNSGSGNTNTTGYPAKYDSVLAV 176
QY 389 GATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTXILSARSSLAPDSSFMA 448
DB 177 GA-----VDSNSNRASFSSVG-----AELVMAPOAGVYSTYPT----- 210
QY 449 NHDSKYAYMGTSMATPIYAGNVA-----QLRHFVKNR 482
DB 211 ---NTYATWDTGTSMASPHVAGAAALILSKHPVLSASQVKNR 248

Search completed: April 7, 2006, 18:04:21
Job time : 18.5129 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:56:02 ; Search time 9406.15 Seconds
(without alignments)
11602.984 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atgagagaagaagaagggtgtt.....cgttgcaatgtgaactaa 1920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hlg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1920	100.0	1920	1	AB046403
2	1920	100.0	1920	6	AR368116
3	1920	100.0	1920	6	AR562464
4	1478.8	77.0	1923	1	AB084155
5	1477.2	76.9	1923	6	CO888285
6	1477.2	76.9	1923	6	CO888285
7	1477.2	76.9	1923	6	CO888285
8	1477.2	76.9	1923	6	AR368117
9	1477.2	76.9	1923	6	AR368118
10	1477.2	76.9	1923	6	AR562465
11	1477.2	76.9	1923	6	AR562466
12	1136.8	59.2	3003	6	BD062155
13	1136.8	59.2	3003	6	BD062155
14	1002.6	52.2	1305	6	BD062155
15	930	42.9	2218	6	E03808
16	823.6	48.4	1302	1	AB046406
17	747.6	38.9	1299	1	AB046404
18	738	38.4	1299	1	AB046405

19	734.8	38.3	1299	1	AB046402	AB046402 Bacillus
20	89	4.6	346148	1	BX842652	BX842652 Bellowib
21	58.2	3.0	7218	6	I66494	I66494 Sequence 14
22	47.4	2.5	2000	6	AX655393	AX655393 Sequence
23	47	2.4	2000	6	AX655393	AX655393 Sequence
24	46	2.4	5772	2	AF466309	AF466309 Dictyoste
25	45.2	2.4	110000	1	BA000004_20	Continuation (21 o
26	44.6	2.3	250541	1	AE017004	Continuation (21 o
27	44.2	2.3	125020	8	AF429315	AE017004 Bacillus
28	44	2.3	250029	2	AE014820	AF429315 Homo sapi
29	43.6	2.3	6115	2	DD020432	AE014820 Plasmodiu
30	43.6	2.3	110000	1	AB017225_18	U20432 Dictyoste
31	43.6	2.3	110000	1	AB017334_18	Continuation (19 o
32	43.6	2.3	110000	1	AB017355_19	Continuation (19 o
33	43.6	2.3	110000	1	CP000001_19	Continuation (20 o
34	43.6	2.3	291490	1	AE017030	Continuation (20 o
35	43.4	2.3	110000	1	AE017180_35	AB017030 Bacillus
36	43.4	2.3	154712	14	AC139734	Continuation (36 o
37	43.4	2.3	211431	14	AC111594	AC139734 Sus scrof
38	43.4	2.3	246852	14	AC096194	AC111594 Ratrus no
39	43.4	2.3	255459	14	AC127743	AC096194 Ratrus no
40	42	2.2	110000	14	TANNA_03	AC127743 Ratrus no
41	41.2	2.1	110000	1	BA000028_24	Continuation (4 of
42	41	2.1	1141	6	AR579680	Continuation (25 o
43	41	2.1	1141	6	AX083744	AR579680 Sequence
44	41	2.1	4839	1	VIBPSA19	AX083744 Sequence
45	41	2.1	168438	14	AC151436	D85144 Vibrio para
						AC151436 Carolinia

ALIGNMENTS

RESULT 1
AB046403
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
11118284
2 (bases 1 to 1920)
Saeki, K.
Direct Submission
Submitted (20-JUN-2000) Katsuhisa Saeki, Kyo Corporation,
Biological Science Laboratory, Ichikaimachi Akabane 2606, Hega,
Tochigi 331-3497, Japan (E-mail:387185@katsanet.kyo.co.jp,
Tel:81-285-68-7400, Fax:81-285-68-7403)
On May 9, 2002 this sequence version replaced gt:12381938.
FEATURES
source
1..1920
/organism="Bacillus sp. 9860"
/mol_type="genomic DNA"
/strain="9860"
/db_xref="taxon:133778"
1..1920
/gene="PROA"
1..1920
/gene="PROA"
/codon_start=1
/transl_table=1
/product="protease"
/protein_id="BAB21266.2"
/db_xref="GI:20521153"

/translation="MRKKVPLSVLAAIITLVLANPNSAGDAPDPLDKIGIOTT
DVSGRSKOROTGAAPFLVSEENVKILKELIKLETVPANNKLIHYQFNPILIBETKOK
LETTAKLIDYI PDIAYI VBYEVDQVSVERSI EHVESYEPILPKIDIPILIGAST
LVKALADTKONNKEVORLGEI EIAQVYASNDVHITKPEKWNDAKRI VADVA
OSSVGLYGOQOI VAAADTGLDTRNDSSMHEFRKILALVAGSTNANNTNGHGT
VAGSVLGAATNKMA POANLVFOSIMSSGGLPBNLOLFSQASASARHITNS
MGAIVMGA VTTDSRVDDYVRKNDTILPAAGNERPNNGTIL SAPETANALTYGTEN
LRPSGAYADININHYAOPSSRQPTKDGKIKPDVA RGYIISASISLAPDSSFANND
SKVAFMGTSMATPIVAGNVAQLREHFVYKNKQITPKPSLKLALAGADVLGTPNG
NQGMRVTLDKSLNVAYNESFASLSOKATYTFATAGKPLKLSLVMSDAPASTTAS
VTLVNDLVLITAPNGRTYVGNDFSAFPDNMDGNHNVENFINSPGTYTIEVOAY
NVPVQPNFSLAIYN"

ORIGIN

Y

Query Match 100.0%; Score 1920; DB 1; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAGTGTCTTTATCTGTCTTATCAGCTGCAGCGATCTGTGACTGTT 60
DB 1 ATGAGAAAGAAAGAGTGTCTTTATCTGTCTTATCAGCTGCAGCGATCTGTGACTGTT 60
QY 61 GCATTAAACAATCCCTCGGCTGTGTATGCAAGAACTTTGATCTGATTTTAAAGAAATT 120
DB 61 GCATTAAACAATCCCTCGGCTGTGTATGCAAGAACTTTGATCTGATTTTAAAGAAATT 120
QY 121 CAAACAAACCCGATGTCAGTGTTCTTCGAAACAGCGACAAACAGTGCCTGTCATT 180
DB 121 CAAACAAACCCGATGTCAGTGTTCTTCGAAACAGCGACAAACAGTGCCTGTCATT 180
QY 181 CTGGTGAAGTCTGAAGATGTGAACCTTCTTAAAGATGTCTTAAAGAACTTGAACAGTA 240
DB 181 CTGGTGAAGTCTGAAGATGTGAACCTTCTTAAAGATGTCTTAAAGAACTTGAACAGTA 240
QY 241 CCGGCAAAATATTAATCTCCATATGTGTCCAATTCAATGCGCCCATTTTAGAAGAAACAAA 300
DB 241 CCGGCAAAATATTAATCTCCATATGTGTCCAATTCAATGCGCCCATTTTAGAAGAAACAAA 300
QY 301 CAGAAAGCTAAGACAACTGAGCAAAAGATCTGCACTATCATCTGATTTATGCAATATTT 360
DB 301 CAGAAAGCTAAGACAACTGAGCAAAAGATCTGCACTATCATCTGATTTATGCAATATTT 360
QY 361 GTGAGTATGAGGGGGATGTTCACTCAAAAGTCCGCTCCATTGAACACGTGGAATCAGTG 420
DB 361 GTGAGTATGAGGGGGATGTTCACTCAAAAGTCCGCTCCATTGAACACGTGGAATCAGTG 420
QY 421 GAGCCATCTTGCAGAAATACAAATATAGATCCCAAGCTTTTCACAAAGGCGCATCGACG 480
DB 421 GAGCCATCTTGCAGAAATACAAATATAGATCCCAAGCTTTTCACAAAGGCGCATCGACG 480
QY 481 CTGGTGAAGCTTGGCGCTTGTATGCAAGGCAAAACAATAAAGATGCAATTAGAGAGCG 540
DB 481 CTGGTGAAGCTTGGCGCTTGTATGCAAGGCAAAACAATAAAGATGCAATTAGAGAGCG 540
QY 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACGTCCATTATATTACGCAAAAGCT 600
DB 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACGTCCATTATATTACGCAAAAGCT 600
QY 601 GAAATTAAGGTATGATATGATGTGGCCAGAGGTAATTTGCAAAAGCGAGTGTGCAAGAGC 660
DB 601 GAAATTAAGGTATGATATGATGTGGCCAGAGGTAATTTGCAAAAGCGAGTGTGCAAGAGC 660
QY 661 AGCTACGGTTGTATGAGCAAGGCCAGATGTGCGAGTTGCCGATACGTGAATTGATTAACA 720
DB 661 AGCTACGGTTGTATGAGCAAGGCCAGATGTGCGAGTTGCCGATACGTGAATTGATTAACA 720
QY 721 GGAAGAAACGACAGTTCAGTACGAGCCTTCGCGGATTAATTAACAGCACTATATATGCA 780
DB 721 GGAAGAAACGACAGTTCAGTACGAGCCTTCGCGGATTAATTAACAGCACTATATATGCA 780
QY 781 CTGGGTCCGACGAATATATGCGAATATACGAAACGGTCATGTATACCCATGTGCAAGGTTCC 840
DB 781 CTGGGTCCGACGAATATATGCGAATATATGCGAATATACGAAACGGTCATGTATACCCATGTGCAAGGTTCC 840

QY 841 GTATTAGAAATGCGCGCAACGAATTAAGGAATGGCACTCTCAAGGAAATCTGGTTTTTCAA 900
DB 841 GTATTAGAAATGCGCGCAACGAATTAAGGAATGGCACTCTCAAGGAAATCTGGTTTTTCAA 900
QY 901 TCCATCATGATGATACAGTGTGTGAGGCTTGTGCAAGCTTTCCTTCAATCTGCAAACTTATTC 960
DB 901 TCCATCATGATGATACAGTGTGTGAGGCTTGTGCAAGCTTTCCTTCAATCTGCAAACTTATTC 960
QY 961 AGCCAGCATTCAGTGCAGGTGCCAGAAATTATACAACTCTCTGGGGGGCAGCGGTGAAT 1020
DB 961 AGCCAGCATTCAGTGCAGGTGCCAGAAATTATACAACTCTCTGGGGGGCAGCGGTGAAT 1020
QY 1021 GGGGCTTACAGACAGATGATCCAGAAATGTGATAGTATATGAGAAATATGATATGACG 1080
DB 1021 GGGGCTTACAGACAGATGATCCAGAAATGTGATAGTATATGAGAAATATGATATGACG 1080
QY 1081 ATTCCTTTCGCGGCTGTGGAATGAAGCGCGAATCCGATCATCATGTCACCTGTATACG 1140
DB 1081 ATTCCTTTCGCGGCTGTGGAATGAAGCGCGAATCCGATCATCATGTCACCTGTATACG 1140
QY 1141 GCTTAAACCGCATTAACAGTGTGCGCAACGAAACCTGCGTCCAGACTTGGTTCCTAT 1200
DB 1141 GCTTAAACCGCATTAACAGTGTGCGCAACGAAACCTGCGTCCAGACTTGGTTCCTAT 1200
QY 1201 GCAGTAAATATTAACAGATTTGCAAGTTCTTCCGCGGCGCAAAAGATGGGCGA 1260
DB 1201 GCAGTAAATATTAACAGATTTGCAAGTTCTTCCGCGGCGCAAAAGATGGGCGA 1260
QY 1261 ATCAAGCTGATGTATGTCATGCGCCAGGACATACATTTTATGAGCAAGATCTTCTTGCA 1320
DB 1261 ATCAAGCTGATGTATGTCATGCGCCAGGACATACATTTTATGAGCAAGATCTTCTTGCA 1320
QY 1321 CCGGATTCCTCTTCTTGGGCGAATCATACAGCAAAATATGCTTATATGGGTGGAACCTCC 1380
DB 1321 CCGGATTCCTCTTCTTGGGCGAATCATACAGCAAAATATGCTTATATGGGTGGAACCTCC 1380
QY 1381 ATGGCAACCCGATTTGTTGCGGGGAATGTTGCAAGCTCCGTGAGCATTTTGTGAAAAT 1440
DB 1381 ATGGCAACCCGATTTGTTGCGGGGAATGTTGCAAGCTCCGTGAGCATTTTGTGAAAAT 1440
QY 1441 AGAGGAATCACTCTTAAGCCTTCCCTATTTGAAGACGCTTGTATGACAGTGTGCTGAT 1500
DB 1441 AGAGGAATCACTCTTAAGCCTTCCCTATTTGAAGACGCTTGTATGACAGTGTGCTGAT 1500
QY 1501 GTTGAATGGGTATCCGAACGGAACCAAGATGGGGCGAGTGAACCTTGATTAATTCG 1560
DB 1501 GTTGAATGGGTATCCGAACGGAACCAAGATGGGGCGAGTGAACCTTGATTAATTCG 1560
QY 1561 TTGAACGTTGCTATGTGAAGAAATCCAGTCCCTATCACTAGCCAAAGACGACATAT 1620
DB 1561 TTGAACGTTGCTATGTGAAGAAATCCAGTCCCTATCACTAGCCAAAGACGACATAT 1620
QY 1621 ACCTTACTGACAGCGCGGCAAGCATTTGAAGAAATCTCTGTGTATGTGCGATGCCCT 1680
DB 1621 ACCTTACTGACAGCGCGGCAAGCATTTGAAGAAATCTCTGTGTATGTGCGATGCCCT 1680
QY 1681 GCAAGCATTAATGCTTCTGTAACTCTGTGTCAATGATTTGGATTTGTGTATACGACACA 1740
DB 1681 GCAAGCATTAATGCTTCTGTAACTCTGTGTCAATGATTTGGATTTGTGTATACGACACA 1740
QY 1741 AACGGAACAATATGTGCGGAATGACTTCTCAGACCAATTTGACAAATATCTGGATGGC 1800
DB 1741 AACGGAACAATATGTGCGGAATGACTTCTCAGACCAATTTGACAAATATCTGGATGGC 1800
QY 1801 CGCAATTAACGTAAGAAATATGATTTATTAATTCGCCCAAGTGAACATATACATTTAG 1860
DB 1801 CGCAATTAACGTAAGAAATATGATTTATTAATTCGCCCAAGTGAACATATACATTTAG 1860
QY 1861 GTGCAAGATTAATATGTCGGTGTGACCAAAAATCTCTGTGGCAATGTGTAACCTTA 1920
DB 1861 GTGCAAGATTAATATGTCGGTGTGACCAAAAATCTCTGTGGCAATGTGTAACCTTA 1920

RESULT 2
AR368116
LOCUS AR368116 1920 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6376227.
ACCESSION AR368116
VERSION AR368116.1 GI:34601777
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Takaiwa, M., Okuda, M., Saeki, K., Kubota, H., Hitomi, J., Kageyama, Y.,
Shikata, S. and Nomura, M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 3 23-APR-2002;
Kao Corporation; Tokyo;
JPK;
FEATURES
source 1. 1920
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1920; DB 6; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAGAGGTTTATCTGTTTATCTGTTTATCTGAGTCAGCGCATCTGCTGT 60
DB 1 ATGAGAAAGAGAGGTTTATCTGTTTATCTGTTTATCTGAGTCAGCGCATCTGCTGT 60
QY 61 GCATTAAACAATCCCTCGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
DB 61 GCATTAAACAATCCCTCGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
QY 61 GATTAAACAATCCCTCGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
DB 61 GATTAAACAATCCCTCGCTGTGTATGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
QY 121 CAACAACAACCGATGTGATGTTTCTCCAAACAGCGCAACAAGGTGCGCTGCAATT 180
DB 121 CAACAACAACCGATGTGATGTTTCTCCAAACAGCGCAACAAGGTGCGCTGCAATT 180
QY 121 CAAACAACAACCGATGTGATGTTTCTCCAAACAGCGCAACAAGGTGCGCTGCAATT 180
DB 121 CAAACAACAACCGATGTGATGTTTCTCCAAACAGCGCAACAAGGTGCGCTGCAATT 180
QY 181 CTGTGAGACTGTAAGAAATGTGAACCTTCTTAAAGATTGCTAAGAACTTGAACAGTA 240
DB 181 CTGTGAGACTGTAAGAAATGTGAACCTTCTTAAAGATTGCTAAGAACTTGAACAGTA 240
QY 181 CTGTGAGACTGTAAGAAATGTGAACCTTCTTAAAGATTGCTAAGAACTTGAACAGTA 240
DB 181 CTGTGAGACTGTAAGAAATGTGAACCTTCTTAAAGATTGCTAAGAACTTGAACAGTA 240
QY 241 CCGGCAAAATAAACTCCATATTTGTCATTCATTCATTCATTCATTCATTCATTCATTC 300
DB 241 CCGGCAAAATAAACTCCATATTTGTCATTCATTCATTCATTCATTCATTCATTCATTC 300
QY 301 CAGAAGCTAGAGCAACTGAGCAAAAGATTCTGACTATCCCTGATTATGATATATT 360
DB 301 CAGAAGCTAGAGCAACTGAGCAAAAGATTCTGACTATCCCTGATTATGATATATT 360
QY 361 GTGAGTATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCCATTTGAACAAGTCAAGT 420
DB 361 GTGAGTATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCCATTTGAACAAGTCAAGT 420
QY 361 GTGAGTATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCCATTTGAACAAGTCAAGT 420
DB 361 GTGAGTATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCCATTTGAACAAGTCAAGT 420
QY 421 GAGCCATATCTGCGCAAAATACAAATAGATCCCAAGCTTTTCAAAAAGGCCATCGAG 480
DB 421 GAGCCATATCTGCGCAAAATACAAATAGATCCCAAGCTTTTCAAAAAGGCCATCGAG 480
QY 481 CTGGTGAAGAGGTTGCGCTGATACGAAGCAAGCAATTAAGAGTGAATTAAGAGGC 540
DB 481 CTGGTGAAGAGGTTGCGCTGATACGAAGCAAGCAATTAAGAGTGAATTAAGAGGC 540
QY 541 ATCGAAGAAATCGCTCAGTACGTAGCAAGCAATGACATCATTTATTTAAGCAAGCCT 600
DB 541 ATCGAAGAAATCGCTCAGTACGTAGCAAGCAATGACATCATTTATTTAAGCAAGCCT 600
QY 601 GAATTAAGAGTATGATGATGTGGCAGAGGATTTTCAAAAGCGGATGCGACAGAGC 660
DB 601 GAATTAAGAGTATGATGATGTGGCAGAGGATTTTCAAAAGCGGATGCGACAGAGC 660
QY 661 AGCTACGGTTTATGAGCAAGGCCAGATGTGCGAGTTGCCGATATCTGATTTGATACA 720

DB 661 AGCTACGGTTTATGAGCAAGGCCAGATGTGCGAGTTGCCGATATCTGATTTGATACA 720
QY 721 GGAAGAAACGACAGTTCGATGCAATGAAAGCTTCCGGGTAAATAACAGCATATATGCA 780
DB 721 GGAAGAAACGACAGTTCGATGCAATGAAAGCTTCCGGGTAAATAACAGCATATATGCA 780
QY 781 CTGGTGGGAGCAATTAATGCGAATGATGATGAAACGGTCAATGATACCATGTGGAGGTTGC 840
DB 781 CTGGTGGGAGCAATTAATGCGAATGATGATGAAACGGTCAATGATACCATGTGGAGGTTGC 840
QY 841 GTATTAGAAATGCGCAACGAATTAAGAAATGAGCACTCAAGCGAATCTGTTTTCAA 900
DB 841 GTATTAGAAATGCGCAACGAATTAAGAAATGAGCACTCAAGCGAATCTGTTTTCAA 900
QY 901 TCCATCATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TCCATCATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGCCAAAGCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1020
DB 961 AGCCAAAGCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1020
QY 1021 GGGGCTTACACAGCAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GGGGCTTACACAGCAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ATTCTTTTCCGCGCTGAGGATGAAAGGCGGAAACGCGGATGATGATGATGATGATGATGAT 1140
DB 1081 ATTCTTTTCCGCGCTGAGGATGAAAGGCGGAAACGCGGATGATGATGATGATGATGATGAT 1140
QY 1141 GCTAAACAAACGCAATTAACAGTGGCGGCAACCGAAACCTGCGTCAAGCTTCGTTCTAT 1200
DB 1141 GCTAAACAAACGCAATTAACAGTGGCGGCAACCGAAACCTGCGTCAAGCTTCGTTCTAT 1200
QY 1201 GGAATTAATTAATTAACAGTGGCGGCAACCGAAACCTGCGTCAAGCTTCGTTCTAT 1260
DB 1201 GGAATTAATTAATTAACAGTGGCGGCAACCGAAACCTGCGTCAAGCTTCGTTCTAT 1260
QY 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 CCGGATTCCTCTTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 CCGGATTCCTCTTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 ATGGCAACCGGATTCCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 ATGGCAACCGGATTCCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 AGAGGAATCACTCTTAAGCTTCCCTAATGAAAGAGCTTTGATTTGACAGTCTGAT 1500
DB 1441 AGAGGAATCACTCTTAAGCTTCCCTAATGAAAGAGCTTTGATTTGACAGTCTGAT 1500
QY 1501 GTTGAATGAGGTTATCCGAACGAAACCAAGGATGAGGCGGAGTCACTGATTAATG 1560
DB 1501 GTTGAATGAGGTTATCCGAACGAAACCAAGGATGAGGCGGAGTCACTGATTAATG 1560
QY 1561 TTGAAGTGTGCTAATGTAAGAAATCCAGTGGCTTCAACATGACCAAAAGCGACATAT 1620
DB 1561 TTGAAGTGTGCTAATGTAAGAAATCCAGTGGCTTCAACATGACCAAAAGCGACATAT 1620
QY 1621 ACCTTTACTGCAACCGGCGGCAAGCACTTGAATAATCTCCCTGATATGATGATGATGATGAT 1680
DB 1621 ACCTTTACTGCAACCGGCGGCAAGCACTTGAATAATCTCCCTGATATGATGATGATGATGAT 1680
QY 1681 GCAAGCACTACTGCTTCTGTAACCTTGTCAATGATTTGATTTGATTTGATTTGATTTGATTT 1740
DB 1681 GCAAGCACTACTGCTTCTGTAACCTTGTCAATGATTTGATTTGATTTGATTTGATTTGATTT 1740
QY 1741 AACGGAACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800

Db 1741 AACGGAACAAGATATGTCCGGAACTTCTCAGACACATTGACATTAACCTGGATGAC 1800
Qy 1801 CGCAATAACGTAAGAAATGTAATTTTATTTATTCGCCCCAAAGTGAACATATACCTTAG 1860
Db 1801 CGCAATAACGTAAGAAATGTAATTTTATTTATTCGCCCCAAAGTGAACATATACCTTAG 1860
Qy 1861 GTGCAAGCATTAATATGTCCGGTGGACCAACAACTTCTCGTTGGCAATGTGAATCA 1920
Db 1861 GTGCAAGCATTAATATGTCCGGTGGACCAACAACTTCTCGTTGGCAATGTGAATCA 1920

RESULT 3
AR562464
LOCUS AR562464 1920 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6759228.
ACCESSION AR562464
VERSION AR562464.1 GI:53976505
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1920)
AUTHORS Takeiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline phosphatase
JOURNAL Patent: US 6759228-A 3 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;

FEATURES
source location/Qualifiers
1..1920
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1920; DB 6; Length 1920;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATCAGCTGCAGCATTCGTCCACTGTT 60
Db 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATCAGCTGCAGCATTCGTCCACTGTT 60
Qy 61 GCATTAAACAATCCCTCGGCTGTGATGACAGGACCTTTTGATCTGGAATTTTAAAGAAAT 120
Db 61 GCATTAAACAATCCCTCGGCTGTGATGACAGGACCTTTTGATCTGGAATTTTAAAGAAAT 120
Qy 121 CAAACCAACACGATGTGATGTTCTCCAAACAGCGACAAACAGTGCAGCTGCATTT 180
Db 121 CAAACCAACACGATGTGATGTTCTCCAAACAGCGACAAACAGTGCAGCTGCATTT 180
Qy 181 CTGGTGAAGTCTGAAGAAATGTGAACTTCTTAAAGATGCTTAAAGAACTTGAACAGTA 240
Db 181 CTGGTGAAGTCTGAAGAAATGTGAACTTCTTAAAGATGCTTAAAGAACTTGAACAGTA 240
Qy 241 CCGGCAATTAATAAATCCCATATTGTCCAAATTCATAGGCCCATTTTGAAGAAACAAA 300
Db 241 CCGGCAATTAATAAATCCCATATTGTCCAAATTCATAGGCCCATTTTGAAGAAACAAA 300
Qy 301 CAGAACTAGAGACAACTGAGCAAAAGATTCGACCTATCCCTGATTAATGATATATTT 360
Db 301 CAGAACTAGAGACAACTGAGCAAAAGATTCGACCTATCCCTGATTAATGATATATTT 360
Qy 361 GTCGAGTATGAGGGGAGATGTTCACTGAAAGTCCGCTCATTTGAACAGTGAATCACTG 420
Db 361 GTCGAGTATGAGGGGAGATGTTCACTGAAAGTCCGCTCATTTGAACAGTGAATCACTG 420
Qy 421 GAGGCATCTCTGCGAAATATCAAAATAGATCCCGAGCTTTTCACAAAGGGGCATCGAG 480
Db 421 GAGGCATCTCTGCGAAATATCAAAATAGATCCCGAGCTTTTCACAAAGGGGCATCGAG 480
Qy 481 CTGGTGAAGAGCGTTGCGCTTGAATCGAAGCAAACTAAAGAGTCAATTAAGAGGC 540
Db 481 CTGGTGAAGAGCGTTGCGCTTGAATCGAAGCAAACTAAAGAGTCAATTAAGAGGC 540

Qy 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACGTCATTAATTAACGGCAAGCCT 600
Db 541 ATCGAGAAATCGCTCAGTACGTAGCAAGCAATGACGTCATTAATTAACGGCAAGCCT 600
Qy 601 GAATATTAAGTGAATGATGTGCGCAGAGTATTTGCAAAAGCGATGTGGCACAGAGC 660
Db 601 GAATATTAAGTGAATGATGTGCGCAGAGTATTTGCAAAAGCGATGTGGCACAGAGC 660
Qy 661 AGCTACGTTTGTATGCAAGAGCCAGATTGTCCGAGTTGCCGATCTGGAATGGAATCA 720
Db 661 AGCTACGTTTGTATGCAAGAGCCAGATTGTCCGAGTTGCCGATCTGGAATGGAATCA 720
Qy 721 GGAAGAAACGACAGTGTGATGAGAGCCCTCCGCGGTAAATTAACGACTATATGCA 780
Db 721 GGAAGAAACGACAGTGTGATGAGAGCCCTCCGCGGTAAATTAACGACTATATGCA 780
Qy 781 CTGGGTGCGAGCAATTAATGCGAAATGATACGAACGGTCACTGTAACCATGTGGCAGTTGC 840
Db 781 CTGGGTGCGAGCAATTAATGCGAAATGATACGAACGGTCACTGTAACCATGTGGCAGTTGC 840
Qy 841 GTATTTAGAAATGCGCGCAAGATTAAGAAATGCGACCTCAAGCGAAATCTGTTTTCAA 900
Db 841 GTATTTAGAAATGCGCGCAAGATTAAGAAATGCGACCTCAAGCGAAATCTGTTTTCAA 900
Qy 901 TCCATCANTGATACAGTGTGCGCTTGAAGGCTTGCCTTCATCTGCAAACTTATTC 960
Db 901 TCCATCANTGATACAGTGTGCGCTTGAAGGCTTGCCTTCATCTGCAAACTTATTC 960
Qy 961 AGCCAAGCATTCAGTGCAGGTGCGCAGAAATTCATCAAACTCTCGGGGGCGACGGTGAAT 1020
Db 961 AGCCAAGCATTCAGTGCAGGTGCGCAGAAATTCATCAAACTCTCGGGGGCGACGGTGAAT 1020
Qy 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATCTATGAGAAATGATATGACG 1080
Db 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATCTATGAGAAATGATATGACG 1080
Qy 1081 ATTCCTTTCGGGCTGGGAAATGAAAGCCGAAACGGCGGTACATCAGTGCACCTGGTACG 1140
Db 1081 ATTCCTTTCGGGCTGGGAAATGAAAGCCGAAACGGCGGTACATCAGTGCACCTGGTACG 1140
Qy 1141 GCTAAAAACGCATTAACAGTGTGCGCAACCGAAACCTGCTCCAAAGCTTGGTTCAT 1200
Db 1141 GCTAAAAACGCATTAACAGTGTGCGCAACCGAAACCTGCTCCAAAGCTTGGTTCAT 1200
Qy 1201 GCAATATATTAACAACGTTGCAAGTCTCTTCCGTTGCGCGACAAAGATGGGCGA 1260
Db 1201 GCAATATATTAACAACGTTGCAAGTCTCTTCCGTTGCGCGACAAAGATGGGCGA 1260
Qy 1261 ATCAAGCTGATGTGATGAGGCGCGGACATACATTTTATCAGAAATCTTCTTGA 1320
Db 1261 ATCAAGCTGATGTGATGAGGCGCGGACATACATTTTATCAGAAATCTTCTTGA 1320
Qy 1321 CCCGATTCCTCTCTGCGGCAATCATGACAGCAATATGCTTATGAGTGAACGTC 1380
Db 1321 CCCGATTCCTCTCTGCGGCAATCATGACAGCAATATGCTTATGAGTGAACGTC 1380
Qy 1381 ATGSCAACACCGATTGTTGCGGGAAATGTTGACAGCTCCGTTGACATTTTGTAAAAAT 1440
Db 1381 ATGSCAACACCGATTGTTGCGGGAAATGTTGACAGCTCCGTTGACATTTTGTAAAAAT 1440
Qy 1441 AGAGAAATCACTCCTAAGCCCTTCCCTATTGAAGACACTTTGATGAGAGTGTGCTGAT 1500
Db 1441 AGAGAAATCACTCCTAAGCCCTTCCCTATTGAAGACACTTTGATGAGAGTGTGCTGAT 1500
Qy 1501 GTTGATTTGGGTTATCCGAAACGAAACCAAGATGGGGCGAGTGAACCTTGATTAATTCG 1560
Db 1501 GTTGATTTGGGTTATCCGAAACGAAACCAAGATGGGGCGAGTGAACCTTGATTAATTCG 1560
Qy 1561 TTGAACGTTGCTATGTGAAGCAATTCAGTGCCTTATCAACTAGCAAAAAGCGACATAT 1620
Db 1561 TTGAACGTTGCTATGTGAAGCAATTCAGTGCCTTATCAACTAGCAAAAAGCGACATAT 1620

QY	1621	ACCTTATCTGCAACGGCCGGGCAAGCATTTGAAATCTCCCTGATATGGTCGGAGTCCCT	1680
Db	1621	ACCTTATCTGCAACGGCCGGGCAAGCATTTGAAATCTCCCTGATATGGTCGGAGTCCCT	1680
QY	1661	GCAGCACTACTGCTTCTGTAAACCTTGGTCAATGAATTTGGTCATTACAGACCA	1740
Db	1661	GCAGCACTACTGCTTCTGTAAACCTTGGTCAATGAATTTGGTCATTACAGACCA	1740
QY	1741	AACGGAAACAGATATGTGCGGGAAATGACCTTTCAGACACATTTGACAAATACCTGGATGGC	1800
Db	1741	AACGGAAACAGATATGTGCGGGAAATGACCTTTCAGACACATTTGACAAATACCTGGATGGC	1800
QY	1801	CGCAATAACGTAGAAATATGTTATTTATTTATTCGCCCCAAAGTGGAAACATATACCATTTGAG	1860
Db	1801	CGCAATAACGTAGAAATATGTTATTTATTTATTCGCCCCAAAGTGGAAACATATACCATTTGAG	1860
QY	1861	GTGCAGAGCATTAATATGTCGGGTGGACCAAAAATCTCTGTTGGCAATTTGTGAATCTTA	1920
Db	1861	GTGCAGAGCATTAATATGTCGGGTGGACCAAAAATCTCTGTTGGCAATTTGTGAATCTTA	1920

RESULT 4			
AB084155			
LOCUS	AB084155	1923 bp	DNA
DEFINITION	Bacillus sp. KSM-9865 gene for protease, complete cds.		
ACCESSION	AB084155		
VERSION	AB084155.1	GI:34392386	

ORGANISM
Bacillus sp. KSM-9865
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus

1 Okuda, M., Saeki, K. and Kobayashi, T.
Bacillus sp. KSM-9665 protease gene
published Only in Database (2003)
2 (baspe 1 to 1923)
Okuda, M., Saeki, K. and Kobayashi, T.
Direct Submission
Submitted (18-APR-2002) Mitsunoshi Okuda, Kao corporation,
Biological Science Laboratory, Ichikaimachi Akabane 2606, Haga
Tchohgi 321-3457, Japan (E-mail: okuda.mitsunoshi@xao.co.jp,
Tel:81-285-68-7543, Fax:81-285-68-7547)

FEATURES	Location/Qualifiers
source	1. .1923

```
/organism="Bacillus sp. KSM-9865"  
/mol_type="genomic DNA"  
/strain="KSM-9865"  
/db_xref="taxon:192495"  
1..1923
```

```

/codon_start=1
/transl_table=11
/product="protease"
/protein_id="BAC82522.1"
/db_xref="GI:34392387"
/translators="MRKCKKXVLSAAAILSTVALSNPAGAGARHNDLDFKIQIOT
TDAGPSQSGGTGAALVYSEBNKKPKGLQKLETVAPNNKGLVQNNGLIETKQ
ELKMGKALIDYIPDAVYIEYSGQKPSATSTEDVESVPTLYRIDPOLPTKMS
ELVAKVADLDTQKQKREQLQGLGIEIQIQAPATSDVLTITAKETKMNVAAGIVKAD
AASSYGLYGGQGGVAVADTLTGDRNDSNHEAPRKTYTLTALERTNNADNVGHGG
HVAGSVYANGTSNGKMAPQANLRFOSIMOSGGGLGGLPENLTQLEPQSVASGAI
HMGAVANGATVSSRNDVYRKNWDTLIPAGNBSGNGTSAGTANALVGAETRE
LILRSPGSGNDINHWAFQSSRPBTDGRIKPDVWAGPFIILNASSILAPOSFMANNH
DSKAVAGGTSMAPIVAGVWADLREHPIKNGIGTPKPSLLKALILADADIGCTPR
GNQCKGRVTLDKSLANAYVBSLSLSQKATYSFPAIAGKGLKLSLWASDAPSTPR
SVTLVNDLIDVLTAPNVSTQVAGNDFTSPNNNNMGNNVENVINAPQSGTYTLEVAQ
INNVVGGQDTLEVALVN"

```

ORIGIN

Query Match	77.0%;	Score 1478.8;	DB 1;	Length 1923;
Best Local Similarity	85.8%;	Pred. No. 0;		
Matches 1642; Conservative	0;	Mismatches 272;	Indels 0;	Gaps 0

QY 7 AAGAGAAGTGTTTTATCTGTTTATCAGCTGCAGCGATTCTGTGACTGTTGCATTA 66

Db	10	AAAAAAAAAGGTTTAACTGTTTATCAGTGTGACGAGATTGTGTCACTGTGGTTA	69
Qy	67	AACATCCCTGGCTGTGATGCAAGACCTTGTGATCTGGAATTTTAAAGAAATTCAAAC	126
Db	70	AGTATTCATCTGCAGGAGTGTGCAAGAAATTTGATCTGGATTTCAAAGAAATTCAGACA	129
Qy	127	ACAAACGATGTCAGTGTGTTCTCCAAACAGCGACAAACAGTGCGGTGCAATTTCTGGTG	186
Db	130	ATAACTGATGCTTAAAGTTTCTCCAAAGACGGGGCAGACTGGGTCTGTGCAATTTCTGGTG	189
Qy	187	GAGCTGTGAAAATGTGAAAATTCTTAAAGATTTGCTAAAGAACTTGAAAACAGTACCGGCA	246
Db	190	GAATCTGAAAATGTGAAAATCTCCCAAAAGTTTGCAGAAAGAGCTTGAAAACAGTCCCGCA	249
Qy	247	AATAATAAATCCCAATATATGTCCAATTTCAATGGCCCCATTTTAAAGAAACAAACAGAAAG	306
Db	250	AATATATAATCTCCATATTTGTCCAAATTCATATGACCAATTTTAAAGAAACAAACAGCAC	309
Qy	307	CTAGAGCAACTGTGAGCAAAAGATTTCTGCATACATCCCTGATTAATATATTTGTGAG	366
Db	310	CTAGAAAAAACAGGGGCAAAGATTTCTGCATACATACCTGATTAATATTTGTGAG	369
Qy	367	TATAGAGGGGATGTTCAATCAAAAGTCCGCTCCATTGAAACAGTGGATCAATGTGAGCCA	426
Db	370	TATAGAGGCGCATGTTAACTAGCAACAAGACCAATTTGAGACGTGTGAATCCGTAAAGCCT	429
Qy	427	TACTTGGCGGAATATCAAAATAGATCCCCAGCTTTTCAAAAAGCGCATTCGACGCTGGTG	486
Db	430	TATTTGCCGATATACAGATATGATATCCAGCTTTTCAAAAAGGGGCATTCAGAGCTTGTA	489
Qy	487	AAAGCGTTGGCGCTTGATATCCAGACAGAACATTAAGAATGCAATTAAGAGGCATTCGAG	546
Db	490	AAAGCAGTGGCGCTTGATATCAAAATCAGAAAAATTAAGAGGTGCAATTAAGAGGCATTCGAG	549
Qy	547	GAATTCGCTCAGTACGTACGACGCAATGACGTCCATTATATTAAGGCAAGCTTGAAATAT	606
Db	550	CAAAATCGCACTATTCGCAACAAGCAATGATGTGTAATATTAACGGCAAAAGCTTGAAATAT	609
Qy	607	AAGGTGATGATGATGATGTGGCCAGAGGTATTTGTCAAAAGCGATGTGTGGCACAGACGCTTAC	666
Db	610	AAGGTGATGATGATGATGTGGCCAGTGGCAATTTGTCAAAAGCGATGTGTGGCCACAGACGCTTAC	669
Qy	667	GGTTTGTATGACAAAGGCCAGATTTGTGCAGTTGCCGATTAATCTGATTTGATATACAGAAAGA	726
Db	670	GGTTTGTATGACAAAGGACAGATGTGTAGCGGTTGCCGATTAACAGGCTTGATACAGGTGGC	729
Qy	727	AACGACAGTTTGATGATGAAAGCCTTCGCGGTAAATTAACAGCACTATATGCACTGGGT	786
Db	730	AATACACAGTTTCGATGATGAAAGCCTTCGCGGGGAAAAATTAATGCACTTATATGCAATGGGA	789
Qy	787	CGGACGAATATATGGAATGATATGCAACGAGTATGATGCCCAATGTGTGGCAGGTTCGGTATTA	846
Db	790	CGGACGAATATATGCAATGATATGCAAAAGTGTATGTATGTATGCAATGTGTGGCTTCGGTATTA	849
Qy	847	GGAATATGGCGCAAGCAATTAAGAATGGCACCTCAAGCAATCTGTTTTCAATCCATC	906
Db	850	GGAATATGGCTCCATTAATTAAGAAATGGCGCTCAAGCGAATCTAGTCTTCCAAATCTATC	909
Qy	907	ATGGAATGACAGTGTGTGGCTTGAGAGCTTGTCTTCCAAATCTGCAAACTTATTCAGCCAA	966
Db	910	ATGGAATGACGGTGTGGGACCTTGAGAGCACTACCTTCGAATCTGCAAACTTATTCAGCCAA	969
Qy	967	GCAATTCAGTGCAGGTGTCAGAAATTCATACAAATCTCGGGGGGACAGCGGTAAATGGGGCC	1022
Db	970	GCAATTCAGTGTGTGTGTCAGAAATTCATACAAATCTCGGGGAGCAAGCAATTAATGGGGCT	1023
Qy	1027	TACACGACAGATTTCCAGAAATGTGTGATGATCTATGTAAAGAAAAATGATATGACGATTTCTT	1088
Db	1030	TACACACAGATTTCCAGAAATGTGTGATGATCTATGTGTGCAAAAAATGATATGACGATCTTT	1089
Qy	1087	TTGCGGCTGTGGAAATGAAAGGCCGACAGCGCGGTATCAATCACTGTGCACTGTGTACGGCTTAA	1146

Db 1090 TTCGCTGCGGGGAATGAAGACCGGACCATCATGACGACGACGACTAA 1149
 Qy 1147 AACCCATTAACAGTGGGCGCAACCGAAACCTGCTGCAAGTTCCTATGACAGAT 1206
 Db 1150 AATGCAATTAACAGTGGGCGCAACCGAAACCTGCTGCAAGTTCCTATGACAGAT 1209
 Qy 1207 AATATTAACAGTGGGCGCAACCGAAACCTGCTGCAAGTTCCTATGACAGAT 1266
 Db 1210 AATATTAACAGTGGGCGCAACCGAAACCTGCTGCAAGTTCCTATGACAGAT 1269
 Qy 1267 CCTGATGTCATGGGCGGACGACATATTTATGACGAAGATCTTCTTTCGACCCGAT 1326
 Db 1270 CCGGATGTCATGGGCGGACGACATATTTATGACGAAGATCTTCTTTCGACCCGAT 1329
 Qy 1327 TCTCTCTTCTGGGCGAATCATGACGAAATATGCTTATGAGTGAACGTCACGCA 1386
 Db 1330 TCTCTCTTCTGGGCGAATCATGACGAAATATGCTTATGAGTGAACGTCACGCA 1389
 Qy 1387 ACACCGATTTGTCGGGGAATGTTGACAGCTCCGTCGACATTTGTTGAAAAATGAGAGA 1446
 Db 1390 ACACCGATTTGTCGGGGAATGTTGACAGCTCCGTCGACATTTGTTGAAAAATGAGAGC 1449
 Qy 1447 ATCACTCTTAAGCCTTCCCTATTTGAAGCAGCTTGAATGACGTCGTCGATGTTGGA 1506
 Db 1450 ATCACTCTTAAGCCTTCCCTATTTGAAGCAGCTTGAATGACGTCGTCGATGTTGGA 1509
 Qy 1507 TTGGTTATCCGAAACGAAACCAAGATGGGCGGAGTGAACCTGGAATTAATCGTTGAAC 1566
 Db 1510 CTGGCTATCCGAAACGATTAACCAAGATGGGCGGAGTGAACCTGGAATTAATCGTTGAAC 1569
 Qy 1567 GTTGCTATTTGGAACGAATTCAGTGCCTTATCAATACCAAAAGCAATATACCTTT 1626
 Db 1570 GTTGCTATTTGGAACGAATTCAGTGCCTTATCAATACCAAAAGCAATATACCTTT 1629
 Qy 1627 ACTGCAACGGGCGGCAACCATTTGAATCTCCCTGATGGTGGATGCCCTGCAAGC 1686
 Db 1630 ACTGCTATGCGCGGCAACCTTTGAAGATCTCCCTGATGGTGGATGCCCTGCAAGC 1689
 Qy 1687 ACTACTGCTTCTGTAACCTTGCTCAATGATTTGATTTGATTAACGACCAAAACGGA 1746
 Db 1690 ACAACTGCTTCTGTAACCTTGCTCAATGATTTGATTTGATTAACGCTCAAAATGGC 1749
 Qy 1747 ACAAGATTTGTCGGGAATGACCTTCACAGCAATTTGACATTAATGAGGAGCGCGAAT 1806
 Db 1750 ACAAGATTTGTCGGGAATGACCTTCACAGCAATTTGACATTAATGAGGAGCGCGAAT 1809
 Qy 1807 AACGTAAGAAATGATTTATTAATCGCCCAAAATGGAACATATACATGAGGTGCA 1866
 Db 1810 AACGTAAGAAATGATTTATTAATCGCCCAAAATGGAACATATGAGGTGCA 1869
 Qy 1867 GCATTAATTAATGTCGGGTTGACACCAAAATCTTCTGTTGGCAATTTGAACTAA 1920
 Db 1870 GCTTATAACGTAACGGGTTGACACCAAGACTTCTGTTGGCAATTTGAACTAA 1923
 RESULT 5
 AB051423 1923 bp DNA linear BCT 10-MAY-2002
 LOCUS Bacillus sp. KP43 PROF gene for protease, complete cds.
 DEFINITION AB051423
 ACCESSION AB051423
 VERSION AB051423.2 GI:20521154
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus sp. KSM-KP43
 Bacillus sp. KSM-KP43
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE
 AUTHORS Itoh, S. and Saeki, K.
 TITLE new protease
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1923)
 AUTHORS Saeki, K.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2000) Katsuhisa Saeki, KAO, CORPORATION,

2606, AKABANE, ICHIKAWACHI, HAGA, TOCHIGI 321-3486, Japan
 (E-mail: 387185@kascanet.kao.co.jp, Tel: 81285687471 (ex. 7471),
 Fax: 81285687403)
 On May 9, 2002 this sequence version replaced gi:14164344.
 COMMENT location/Qualifiers
 FEATURES
 SOURCE
 1..1923
 /organism="Bacillus sp. KSM-KP43"
 /mol_type="genomic DNA"
 /strain="KP43"
 /db_xref="taxon:109322"
 1..1923
 /gene="PRO"
 1..1923
 /gene="PRO"
 /codon_start=1
 /transl_table=1
 /product="protease"
 /protein_id="BAB55674.2"
 /db_xref="GI:20521155"
 /translation="MRKKKVFPLSVSAAAILSTVALSNPSAGARFPDLPKIQTT
 TDAKFSKQQTGAALFLVSEENVKLPKGLQKLETPANNKLIHQNPETLETKQ
 QLEKTKAKLIDYIPDYAVEIBEDVNSATSTIEHVSVEPYLPYRIDQLFTKGS
 ELVAVALDTRKOKKEVOLRIGBQAPALSNVDLYTAKPEYKMNDAVARGVKA
 AQSRYGLYGGQIVAVADTGLDTRNDSNHEAPRKITLVALGRNNANDTGHGT
 HVGSVYGNSTKKGMAPOANIVPOSIMDSGGIGGIPSNLOTLPSQAVAGARIH
 SWGANVAGATTTDSRNDYVRKNDITILPAAGBEPNGCTISAPGTAKALITGATE
 NLRPSFSDVADNINHVAFSSRPTKGRIRKPDMAVGITILSRSSLDPSFWAN
 DSKYAVNGGSMAPPIVAGNVADLRHFVNRGTTPRPSLKALIGADIGVYN
 GNOGMDGVTTDKSLNVAVNVESSLSQKATYSFTYATKPEKLSIWSVAPASST
 SVTLNDLDELVTAPNGTGYVNDFTSPYNDMDGRNNEVFIINARQSGTYTIEVQA
 YNPVVGQFTSLAIIVN"
 ORIGIN
 Query Match 76.9%; Score 1477.2; DB 1; Length 1923;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
 Qy 7 AAGAAAGAGTGTGTTTATGCTGTTTATGACCTGACGAGATTCGTGACCTGTCATTA 66
 Db 10 AAGAAAGAGTGTGTTTATGCTGTTTATGACCTGACGAGATTCGTGACCTGTCATTA 69
 Qy 67 AACATCCCTGGCTGGTGAAGGACTTTGATCGATTTGATTTAAAGAAATCAACA 126
 Db 70 AGTATCATTTGACAGGTGGTGAAGAAATTTGATCTGATTTCAAGAAATTCAGACA 129
 Qy 127 ACAACGATGTCAGTGTGTTCTCCAAACAGGCAACAAAGGTGGGCTGCAATTTCTGGTG 186
 Db 130 ACACTGATGTAAAGGTTTCTCAAGCAGGGGAGAGCTGGTGTGCTGCTTTCTGGTG 189
 Qy 187 GAGTCTGAATAATGTAACCTCTTAAGGATTCCTTAAGAAACCTTGAACAGTACCGCA 246
 Db 190 GAACTGAAATATGTAACCTCCAAAGGTTTGAAGAAAGCTTGAACAGTCCCGCA 249
 Qy 247 AATATATACTCATATTTGTCATATGACGAGCCCATTTTGAAGAAACAAACAGAAAG 306
 Db 250 AATATATACTCATATTTGTCATATGACGAGCCCATTTTGAAGAAACAAACAGAAAG 309
 Qy 307 CTAGAGCAACTGAGCAAGAAATTCGACATACATCCCTGATTAATGATATATTTGCGAG 366
 Db 310 CTGAAAAAACAGGCGGCAAGATTCGACATACATCTGATTAATGCTTACATTTGCGAG 369
 Qy 367 TATGAGGGGATGTTCTGACGTAAGAGTCCGCTCATTTGAACACGCTGGAATCAGTGAACCA 426
 Db 370 TATGAGGGGATGTTGATGTCGACAAACGACCATTTGAGCAGCTGGAATTCGTTGAGGCT 429
 Qy 427 TACTTGCGAAATCAAAATGATCCGAGCTTTTCAAAAGGCGCATGCAAGCTGGTG 486
 Db 430 TATTGCGATATACAGATGATCCCGAGCTTTTCAAAAGGCGCATGCAAGCTTTGTA 489
 Qy 487 AAACGTTGGGCTTGTATACGACAGCAACATTAAGAGTCAATTAAGAGCATGAG 546
 Db 490 AAACGAGTGGGCTTGTATACAAACAGAAATTAAGAGGTCATTAAGAGCATGAG 549

127 ACAACCGATGTCAGTGGTTTCTCCAAACAGGCAAAAGTGGCTGCATTTCTGGTG 186
 130 ACACATGATGCTAAAGGTTTCTCCAGGCAAGGCGGAGATGCTGCTGCTTTCTGGTG 189
 187 GAGTCTGAAAATGTAACCTTCTTAAAGGATGCTTAAAGAACTTGAACAGTACCGCA 246
 190 GATTTGAAAATGTAACCTTCCAAAAGTTTGCAGAAAGCTTGAACAGTCCCGCA 249
 247 AATAATTAATCTCCATATTTGTCATTCATTCAGTGGCTCCCATTTTGAAGAAACAAACAGAG 306
 250 AATAATTAATCTCCATATTTGTCATTCATTCAGTGGCTCCCATTTTGAAGAAACAAACAGAG 309
 307 CTAGAGCAACTGAGGCAAAAGTTTCTCCAGTACATCCCTGATTTATGCAATATTTCTGGAG 366
 310 CTGAAAAAACAAGGCGCAAAAGTTTCTCCAGTACATTCAGTGGCTTCTTACATTTCTGGAG 369
 367 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCCATTTGAACAGTGGAAATCAGTGGAGCA 426
 370 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCCATTTGAACAGTGGAAATCAGTGGAGCT 429
 427 TACTTGGCGAAATACAAATATAGATCCCAAGCTTTTCAAAAAGCGCATTCAGCTGGTG 486
 430 TATTTGGCGAAATACAAATATAGATCCCAAGCTTTTCAAAAAGCGCATTCAGCTGGTG 489
 487 AAAAGCTGGCGCTGATTCAGAGCAAAATTAAGAAAGTGCATTTAAGAGGCAATTCAG 546
 490 AAAAGCTGGCGCTGATTCAGAGCAAAATTAAGAAAGTGCATTTAAGAGGCAATTCAG 549
 547 GAAATCGCTCAGTACGTAGCAAGCAATGACCTTATATATTCAGGCAAAAGCTGAAATAT 606
 550 CAATTCGCAATTCGTAGCAATGAGATGATGCTATATATTAAGGCAAAAGCTGAAATAT 609
 607 AAGGTATGAAATGATGTGGCCAGAGGTATTTGCAAAAGGAGTGTGGCAACAGAGCTTAC 666
 610 AAGGTATGAAATGATGTGGCCAGAGGTATTTGCAAAAGGAGTGTGGCAACAGAGCTTAC 669
 667 GATTGTATGCAAGGCGCAATTTGTCAGTGGCGGATTCAGATTTGGAATTCAGAGAA 726
 670 GGGTTTATGCAAGGCGCAATTTGTCAGTGGCGGATTCAGATTTGGAATTCAGAGAA 729
 727 AACGACAGTTCAGTTCAGTAAAGCTTCCGCGGTAAATTAACAGCACTATATGCACTGGGT 786
 730 AATGACAGTTCAGTTCAGTAAAGCTTCCGCGGTAAATTAACAGCACTATATGCACTGGGT 789
 787 CGGACGAAATTAATGCGAAATGATTCAGAGGCTCATGTTCCATGTGGCGGTTGGTATTA 846
 790 CGGACGAAATTAATGCGAAATGATTCAGAGGCTCATGTTCCATGTGGCGGTTGGTATTA 849
 847 GAAATTCGCGCAAGATTAAGAAATGCGAAGCTTCAAGGCAATTCGTTTTCATTCATCCAT 906
 850 GAAATTCGCGCAAGATTAAGAAATGCGAAGCTTCAAGGCAATTCGTTTTCATTCATCCAT 909
 907 ATGATTCAGTTCAGTTCAGTAAAGCTTCCGCGGTAAATTAACAGCACTATATGCACTGGGT 966
 910 ATGATTCAGTTCAGTTCAGTAAAGCTTCCGCGGTAAATTAACAGCACTATATGCACTGGGT 969
 967 GCATTTCAGTTCAGTTCAGTAAATTCATCAAACTTCTGGGGGCGAGCGGTGATTTGGGGCC 1026
 970 GCATTTCAGTTCAGTTCAGTAAATTCATCAAACTTCTGGGGGCGAGCGGTGATTTGGGGCC 1029
 1027 TACACGACAGATTCAGTTCAGTAAATTCATTCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1086
 1030 TACACGACAGATTCAGTTCAGTAAATTCATTCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1089
 1087 TTGCGGCTGGGAAATGAAGGCGCAACGCGGTACCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1146
 1090 TTGCGGCTGGGAAATGAAGGCGCAACGCGGTACCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1149
 1147 AACGCAATTAACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1206
 1150 AATGCAATTAACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1209

QY 1207 AATAATTAACGAGTTCGACAGTTCCTTCCGTTGGCCCGCAAAAGATGGGCAATCAG 1266
 Db 1210 AATATCAACATGATGAGCAAGTTCCTTCACTGAGGACGCAAAAGATGGGCAATCAG 1269
 QY 1267 CCTGATGTCATGAGGCGCAGGAGCATATTTTATCAGAGATCTTCTTTCAGCCGAT 1326
 Db 1270 CCGATGTCATGAGGCAAGGAGATCTTATCTATCAGAGATCTTCTTTCAGCCGAT 1329
 QY 1327 TCTCTCTTGGGGAATCATGACAGCAATATGCTATATGAGTGGAAAGCTCCATGGCA 1386
 Db 1330 TCTCTCTTGGGGAATCATGACAGCAATATGCTATATGAGTGGAAAGCTCCATGGCT 1389
 QY 1387 ACACGATGTTGGGGAATGTTGACAGCTCCGTTGAGCAATTTGGAATAATGAGGA 1446
 Db 1390 ACACGATGTTGGGGAATGTTGACAGCTCCGTTGAGCAATTTGGAATAATGAGGA 1449
 QY 1447 ATCACTCCTAAGCTTCCCTATTTGAAGAGCTTGAATTCAGGAGTGTGATGTTGGA 1506
 Db 1450 ATCACTCCTAAGCTTCCCTATTTGAAGAGCTTGAATTCAGGAGTGTGATGTTGGA 1509
 QY 1507 TTGGGTTATCCGAACGGAACCAAGATGGGCGAGTACCTTGATTAATTCGTTGAC 1566
 Db 1510 CTGGGTTATCCGAACGGAACCAAGATGGGCGAGTACCTTGATTAATTCGTTGAC 1569
 QY 1567 GTTGCTATGTAAGCAATCCAGTCCCTATCACTAGCCCAAAAGGCAATATACCTTT 1626
 Db 1570 GTTGCTATGTAAGCAATCCAGTCCCTATCACTAGCCCAAAAGGCAATATACCTTT 1629
 QY 1627 ACTGCAACGCGGCGCAAGCCATTGAAATCTCCTGTGATGTTGATGCTCCGCAAGC 1686
 Db 1630 ACTGCAACGCGGCGCAAGCCATTGAAATCTCCTGTGATGTTGATGCTCCGCAAGC 1689
 QY 1687 ACTACTGCTTGTGTAACCTGCTGATGATTTGATTTGATTCATTCAGCAACCAAGGA 1746
 Db 1690 ACAACTGCTTGTGTAACCTGCTGATGATTTGATTTGATTCATTCAGCAACCAAGGA 1749
 QY 1747 ACAAGATATGTCGGGAAATGATCTTCCAGCACTTTGCAATTAAGTGGAGTGGCGCAT 1806
 Db 1750 ACAAGATATGTCGGGAAATGATCTTCCAGCACTTTGCAATTAAGTGGAGTGGCGCAT 1809
 QY 1807 AACGTAGAAATGATTTATTTAATTTGCCCCCAAGTGGAAATATACATTTAGTGGCA 1866
 Db 1810 AACGTAGAAATGATTTATTTAATTTGCAACCAAGGCGGAGTATTAATTAAGTGGCA 1869
 QY 1867 GCATTAATGTCGGGTTGACCAAACTTCTGTTGGCAATTTGGAATTA 1920
 Db 1870 GCTTATTAACGTAACGTTTGAACCAAGACCTTCTGTTGGCAATTTGGAATTA 1923

RESULT 7
 CO889309 1923 bp DNA linear PAT 19-OCT-2004
 LOCUS CO889309
 DEFINITION Sequence 2 from Patent EPI466962.
 ACCESSION CO889309
 VERSION CO889309.1 GI:54305176
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus sp. KSM-KP43
 Bacillus sp. KSM-KP43
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE
 1 Okuda,M., Sato,T., Takimura,Y., Sumitomo,N. and Kobayashi,T.
 Alkaline protease
 JOURNAL Patent: EP 1466962-A 2 13-OCT-2004;
 Kao Corporation (JP)
 FEATURES
 source
 1. 1923
 /organism="Bacillus sp. KSM-KP43"
 /mol_type="unassigned DNA"
 /db_xref="taxon:109322"
 1. 1923
 /note="unassigned protein product"
 /codon_start=1
 /transl_table=1

Y:

```
/protein_id="CAH6186.1"
/db_xref="GI:54305177"
/translation="MRKKKVLIVLSAAIISTVLAISNPAGNRPDLPKGIQTT
TDKAPSKGGQTGAALFVBSENVKLPKIQKELTPVANKLHIIQFNPILSETKQ
ELKTKGAKLIDYIPYAVIYBYEGDVKSATSTIBHVSVEPLPYRIDPOLFKGAS
BIKXVALDITKQNKREVLQJLQIAQFASINDVLYITKPEPKWMDVARIIVADV
AOSSYLQCOQIIVAVDGLDTPGRNDSMEHAFRQKLTALPALRTNNANNGHGT
HVASVLGNGSTNKGAPQANI.VPOSINDSGGLGIPENLTLPQOAYSAARITHN
SWGAAYNGAYTTDSNRVDVYAKNDMTILPAAGNRPNGGTISAGTAKNATTVGATE
NIRPSFGSTADININVAQPSRQPTKDGIKPDVMAPGYFIIASNSISLAPSWANH
DSKYVMGTSMAPIIVAGNVSAQLEHFPVKNGLTPKPSLPAALJAGAADIGLYPN
GKQGWGRVTLDSKLNVAYVNBSSLSQSKATYSFLATGKPLKLSLWSDAPASTA
SVTLVNDLDELVTAPNGTQYGVNDFTSPYNDMDGNVNFVINAPOSQYTTIVQA
YVNPVGPOTFSLAIYN"
1.618
619..1920
/product="unnamed"
```

ORIGIN

```
Query Match      76.9%; Score 1477.2; DB 6; Length 1923;
Beet Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
```

```
QY 7 AAGAGAAAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTCTGCACTGTTGATTA 66
DB 10 AAGAAAAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTGTCGACTGTTGCTTA 69
QY 67 AACAAATCCCTCGGCTGGTGATGCAAGAACTTTTGATCTGGATTTTAAAGAAATTCAAACA 126
DB 70 AGTATTCATCTGCAGTGTGTCAAGAAATTTTGATCTGGATTTTAAAGAAATTCAGACA 129
QY 127 ACAACGATGTGAGTGTGTTTCTCCAAACGCGACAAACGATGCGCGTGCATTTCTGCGG 186
DB 130 ACAACTGATGCTAAAGTGTCTCCAAACGCGGCGAGACTGATGCTGCTGCTTTCTGCGG 189
QY 187 GAGTCTGAATAATGTGAACCTTCTTAAAGATGTCTTAAAGAACTTGAACAGTACCGGCA 246
DB 190 GAATCTGAATAATGTGAACCTTCCCAAGGTTTGCAAGAGAACCTTGAACAGTCCCGGCA 249
QY 247 AATAATTAATCTCATATTTGTCCAAATTCATATGAGCCCCCATTTTAAAGAAACAAACAGAA 306
DB 250 AATAATTAATCTCATATTTGTCCAAATTCATATGAGCCAAATTTTAAAGAAACAAACAGCAG 309
QY 307 CTAGAGACAACTGAGGCAAAAGATTTCTGACATATCTCCGATTTATGCAATATTTGTCCAG 366
DB 310 CTGAGAAAACAGGGGCAAAAGATTTCTGACATATCTGATTTATGCTTACATTTGTCCAG 369
QY 367 TATGAGGGGAGATGTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGAAGCA 426
DB 370 TATGAGGGGAGATGTTCAGTCAAGCAACAGACATTTGAGCAGTGAATTCGTTGAAGCTT 429
QY 427 TACTTCCGGAATTAACAAATAGATCCCAAGCTTTTCACAAAGGCGCATGACGCTGATG 486
DB 430 TATTTGCCGATATACAGAAATAGATCCCAAGCTTTTCACAAAGGCGCATGACGCTTGA 489
QY 487 AAAAGCTTGGCGCTTGTATCGAAGCAACATTAAGAAAGTGCATTTAAGAGGCATCGAG 546
DB 490 AAAAGAGTGGCGCTTGTATCGAAGCAACAAATAAGAGTGCATTTAAGAGGCATCGAA 549
QY 547 GAATTCGCTCAGTACGTAGCAAGCAATGATCCCATTAATTAACGCAAAAGCTGAATAT 606
DB 550 CAATTCGCAATTTGCAATAGCAATGATGCTATATATTAACGCAAAAGCTTGATAT 609
QY 607 AAGGTATGAATGATGTGGCCAGAGGTATTTCAAGCGGATGTGGCAACAGACAGTAC 666
DB 610 AAGGTATGAATGATGTGGCGGGAATTTCAAGCGGATGTGGCTCAGAGCAGCTAC 669
QY 667 GGTGTGTATGACAAAGGCGCAATTTGTCCGATTTGCCGATCTGCAATTTGATACAGAGAA 726
DB 670 GGTGTGTATGACAAAGGCAAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTCC 729
QY 727 AACGACGTTGATGATGATAGCTTCCGCGGTAAATAAGCACTATATGACATGAGT 786
```

```
DB 730 AATGACAGTTCGATGATGATAGAGCTTCCGCGGGAAAAATTAATGCAATTATATGCAATTGGGA 789
QY 781 CGAGGAATTAATGCGAAATGATACGAACGCTCATGCTACCCATGTGGCAGGTTCCGATATTA 846
DB 790 CGAGGAATTAATGCGAAATGATACGAATGCTCATGTGTACGCAATGTGGCTCCGATATTA 849
QY 847 GGAATATGGCGCAACGAATTAAGAAATGGCACTCAAGCGAAATCTGATTTTCAATCATTC 906
DB 850 GGAATATGGCTTCACTAATTAAGAAATGGCGCTCAAGCGAAATCTGATTTTCAATCATTC 909
QY 907 ATGATATGACAGTGTGGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATATTCAGCCA 966
DB 910 ATGATATGACAGTGTGGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATATTCAGCCA 969
QY 967 GCATTCAGTGCAGGTGCGCAAAATTCATACAACTCTTGGGGGCGACCGGTGAATGGGCGC 1026
DB 970 GCATTCAGTGCAGGTGCGCAAAATTCATACAACTCTTGGGGGCGACAGTGAATGGGCGC 1029
QY 1027 TACACGACGATTTCCAGAAATGTGATGATGTAAGGAAAAATGATATGACGATTTCTT 1086
DB 1030 TACACGACGATTTCCAGAAATGTGATGATGTAAGGAAAAATGATATGACGATTTCTT 1089
QY 1087 TTGCGGCGCTGGGAATGAAGGCGGACCGGATCATCATGATGCACTTGTACGCTTAA 1146
DB 1090 TTGCGGCGCGGAATGAAGGACCGAACCGGCGGAACCATCATGATGCACTGAGGCAAGCTTAA 1149
QY 1147 AACGCAATACAGTGTGGCGCAACCGAAACCTGCGCTCAAAGCTTGGGTTCTATGCAAT 1206
DB 1150 AATGCAATACAGTGTGGAGCTTACGGAACCTCCCGCCAAAGCTTGGGTTCTATGCGGAC 1209
QY 1207 AATAATTAACAGTGTGCAAGTGTCTTCCGATGCGCGACAAAGATGGGGAATCAAG 1266
DB 1210 AATAATTAACAGTGTGCAAGTGTCTTCTTCAAGTGAACGACAAAGATGACGATCAAA 1269
QY 1267 CCTGATGTATGAGCCGCGGAGGACATATTTATGACGAAGATCTTCTTCCACCGAT 1326
DB 1270 CCGATGTATGATGAGCAGCGGAACGTTCAATCATGAGCAAGATCTTCTTCCACCGAT 1329
QY 1327 TCTCTCTTCTGGGGCAATCATGACGAATATGCTTATGAGTGAAGGTTCCATGCA 1386
DB 1330 TCTCTCTTCTGGGGCAACATGACAGTAAATATGATCATATGAGTGAAGGTTCCATGCA 1389
QY 1387 ACACGATTTGTCGGGGAAATGTGACAGCTCCGATGACATTTGTGAAAAATGAGGA 1446
DB 1390 ACACGATTTGTCGGGAAACGTGACAGCTTCTGAGCAATTTGTGAAAAATGAGGA 1449
QY 1447 ATCACTCTTAAGCTTCCCTATTTGAAGCAGCTTTGATTTGAGTGTGCTGATGTTGA 1506
DB 1450 ATCACTCTTAAGCTTCTCTATTAAGCAGCACTGATTTGCCGATGAGCTGACATCGG 1509
QY 1507 TTGGGTTATCCGAAACGAAACCAAGATGGGCGGAGTGAACCTTGATTAATGTTTGAAC 1566
DB 1510 CTGGCTACCGGAACGATTAACCAAGATGGGGAACAGTGAACATTTGATTAATCCCTGAAC 1569
QY 1567 GTTGGCTATGTAAGCAATTCAGATGCTTATCAATACACCAAAAGCGACATATACCTTT 1626
DB 1570 GTTGGCTATGTAAGCAATTCAGATGCTTATCAATACCAAAAGCGACATATACCTTT 1629
QY 1627 ACTGCAACGGGCGGCAACGATTAAGAAATCTCCCTGATGATGGTGGATGAGCCCTGCAAC 1686
DB 1630 ACTGCTATCGCGGCAACGCTTTGAAGATCTCCCTGATGATGGTGGATGAGCCCTGCAAC 1689
QY 1687 ACTACTGCTTCTGTAACCTTGTGCAATGATTTGATTTGCTATTAACAGCAACGAACGA 1746
DB 1690 ACAACTGCTTCCGTAACGCTTGTCAATGATCTGAGACCTTGTCACTACGCTCAATATGAC 1749
QY 1747 ACAAGATATGTCGGGAATGATCTTCAGACCACTTTGAACAATAATGGAGTGGCGCAAT 1806
DB 1750 ACAAGATATGTAAGGAATGATCTTTCGCAATCAATGATTAACCTGAGATGGCGCAAT 1809
QY 1807 AACGTAGAAATATGATTTATTAATTCGCCCCAAGTGAACATATACATTTGAGGTGCA 1866
DB 1810 AACGTAGAAATATGATTTATTAATTCACACAAAGCGGAGCGTATACATTTGAGGTACAG 1869
```

Query 1867 GCATTAATGTCGCCGTTGACCAAGAACTTCTGTTGGCAATTGTGACTAA 1920
Db 1870 GCTTATTAACGTACCGGTTGACCAAGACCTTCTGTTGGCAATTGTGACTAA 1923

RESULT 8
AR368117
LOCUS AR368117 1923 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 5 from patent US 6376227.
ACCESSION AR368117
VERSION AR368117.1 GI:34601778
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 5 23-APR-2002;
Kao Corporation; Tokyo;
JPX;

FEATURES
SOURCE 1. 1923
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 76.9%; Score 1477.2; DB 6; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Query 7 AAGAAAGAGTGTTTTATCTGTTTATCACTGACGAGCGATTCTGTGACTTGA 66
Db 10 AAGAAAGAGTGTTTTATCTGTTTATCACTGACGAGCGATTCTGTGACTTGA 69
Query 67 AACATCCCTGGCGTGTGATGACCAAGACTTTGATCTGTGATTTAAAGAAATTCACA 126
Db 70 AGTAATCATCTGACAGGTGTGACCAAGAAATTTGATCTGTGATTTAAAGAAATTCACA 129
Query 127 ACAACGATGTCAAGTGTCTCCAAACAGCGCAAAAGAGTGGCGTCAATTTCTGTG 186
Db 130 ACAATGATGTCAAGTGTCTCCAAACAGCGCGAAGCTGTGTCTGTCTTTCTGTG 189
Query 187 GAGTCTGAAAATGTGAACTTTCTTAAAGGATTTGCTAAAGAACTTGAACAGTACCGCA 246
Db 190 GAACTCTGAAAATGTGAACTCTCCAAAGGTTTGCAGAAAGAGCTTGAACAGTCCGCA 249
Query 247 AATAATAAATCCCAATATGTCCAAATTCATGAGCCCATTTTGAAGAAACAAACAGAG 306
Db 250 AATAATAAATCCCAATATGTCCAAATTCATGAGCCCAATTTTGAAGAAACAAACAGAG 309
Query 307 CTAGAGCAACTGAGCAAGAAATTTCTGACTACATCCCTGATATATGATATTTGAG 366
Db 310 CTGAGAAACAAAGGCGCAAAAGTTCTCGACTACATCACTGATATGCTTACATTTGAG 369
Query 367 TATAGAGGAGATGTTCAGTCAAAAGTCCGCTCCATTGAACAGTGAATCAGTGAAGCA 426
Db 370 TATAGAGGAGATGTTCAGTCAAAAGTCCGCTCCATTGAACAGTGAATCAGTGAAGCA 429
Query 427 TACTTGGCGAATCAAAATATGATCCCGAGCTTTTCAAAAGGCGCATCGAGTGTG 486
Db 430 TATTTGGCGAATCAAAATATGATCCCGAGCTTTTCAAAAGGCGCATCGAGTGTGTA 489
Query 487 AAACGTTGGCGCTTGTATACAGAGCAAGAAATTAAGAGTGAATTAAGAGCATCGAG 546
Db 490 AAACGAGTGGCGCTTGTATACAGAGCAAGAAATTAAGAGTGAATTAAGAGCATCGAA 549
Query 547 GAAATCGCTCAAGTACGTAAGCAAGCAATGATCTTATATTAACGCAAGCTGGAATAT 606
Db 550 CAAATCGCAATTCGCAATTAAGCAATGATGTGCTATATATTAACGCAAGCTGGAATAT 609

Query 607 AAGGTATGATATGATGTGCGCAGAGTATTTGTCAAGCGGATGTGGCAGAGAGCATAC 666
Db 610 AAGGTATGATATGATGTGCGCGTGAATTTGTCAAGCGGATGTGGCTCAGAGAGCATAC 669
Query 667 GGTTTGATATGACAGAGCCAGATTTGTGCAAGTTCGCCATCTGTGATGTGATACAGAGAGA 726
Db 670 GGTTTGATATGACAGAGCAGATGTGAGCGGTTTCCCATATCAGAGGCTTGTATACAGGTGCG 729
Query 727 AACGACAGTTGATGATGAAGCTTCCGCGGTAATAATACAGACATATATGACTGCGGT 786
Db 730 AATGACAGTTGATGATGAAGCTTCCGCGGTAATAATACAGACATATATGACTGCGGA 789
Query 787 CGGACGAATAATGGAATGATATGACAGAGTCAATGCTACATGAGAGGTTCCGTAATTA 846
Db 790 CGGACGAATAATGCAATGATATGACAGAGTCAATGCTACATGAGAGGTTCCGTAATTA 849
Query 847 GGAATGGCGCGCAAGAAATTAAGAAATGAGCACTTCAGACGATCTGTTTTTCAATTCATC 906
Db 850 GGAATGGCGCTCACTAATTAAGAAATGAGCACTTCAGAGGATCTAAGCTTTCATATCATC 909
Query 907 ATGATATGACAGTGTGCGCTTGAAGGCTTGCCTTCAATCTGCAACCTTATTAAGCCAA 966
Db 910 ATGATATGCGGTGGGAGCTTGAAGGACTTCTGAACTGCAAACTTATTAAGCCAA 969
Query 967 GCATTCAGTGAAGGTGCAAGAAATTCATCAAACTCTGGGGGCGAGCGGTGAATGGGGCC 1026
Db 970 GCATTCAGTGTGTGTGCAAGAAATTCATCAAACTCTGGGGGAGGAGGATGAATGGGGCT 1029
Query 1027 TACACGACAGATTCAGAAATGTGATGATGATATGAAGAAATGATATGACAGATTTCTT 1086
Db 1030 TACACACAGATTCAGAAATGTGATGATGATATGAGGCAAAATGATATGACAGATTTCTT 1089
Query 1087 TTGGCGCTGGAAATGAAGCGCGAAGCGGAGTCAATCAAGTCACTGTGACGCTGAA 1146
Db 1090 TTGGCTCGCGGAATGAAGCGCGAAGCGGAGTCAATCAAGTCACTGTGACGCTGAA 1149
Query 1147 AAGGCATTAACAGTGGGCGCAACGGAACCTGTGCTCAAGCTTGGTTCTATGACAGAT 1206
Db 1150 AATGCAATTAACAGTGGAGCTTACGGAACCTCGCCCAAGCTTGGGTCTTATGCGGAC 1209
Query 1207 AATATTAACAGCTTGAACAGTCTCTCCCGTGGGCCCGCAAAAGATGGCGAATGACAG 1266
Db 1210 AATATCAACAGTGTGAGCAAGTCTCTTCAAGTGAACGCAAAAGATGAGCGATCAAA 1269
Query 1267 CTTGATGTCAATGGCGCAAGGACATCAATTTATCAAGAGATTTCTTTCACCCGAT 1326
Db 1270 CCGGATGTCAATGGCGCAAGGACCTTCAATCAATCAAGAGATTTCTTTCACCCGAT 1329
Query 1327 TCCCTCTTGGGCGAATCAATGACAGCAATATGCTATATGGGTGAAGCTTCATGGCA 1386
Db 1330 TCCCTCTTGGGCGAATCAATGACAGCAATATGCTATATGGGTGAAGCTTCATGGCT 1389
Query 1387 ACAACGATTTGTGGGCGAATGTGCAAGCTCCGTGAGCATTTTGTGAATAATAGAGA 1446
Db 1390 ACAACGATTTGTGTGGAAGTGTGCAAGCTTGTGAGCATTTTGTGAATAATAGAGG 1449
Query 1447 ATCACTCTTAAGCTTCTCTATTAAGACAGCTTTGATTTGAGAGTGTCTGTGATTTGA 1506
Db 1450 ATCACTCTTAAGCTTCTCTATTAAGACAGCTTTGATTTGAGAGTGTGTGATGTGCG 1509
Query 1507 TTGGGTTATCGGAACGGAACCAAGATGGGGCGAGTGAACCTCGATTAATGCTTGAC 1566
Db 1510 TTGGGTTATCGGAACGGAACCAAGATGGGGCGAGTGAACCTCGATTAATGCTTGAC 1569
Query 1567 GTTGCTATGTGAACGAATCAAGTGCCTTATCAACTAGCCAAAGGACATATATCTTT 1626
Db 1570 GTTGCTATGTGAACGAATCAAGTGCCTTATCAACTAGCCAAAGGACATATATCTTT 1629
Query 1627 ACTGCAAGCGGCGCAAGCATTTGAATAATCTCCTGTATGTGTGATGCCCTGCAAGC 1686
Db 1630 ACTGCTATCTCGCGCAAGCCTTTGAATAATCTCCTGTATGTGTGATGCCCTGCAAGC 1689
Query 1687 ACTACTGCTTCTGTAAACCTGTGTGAATGATTTGATTTGATCAAGCACCAAGGGA 1746

Db 1690 ACAACGCTCCCTTAACGCTGTCATGATCGAACCTTGTCTCATTCACGCTCCCAATGCG 1749
Qy 1747 ACAAGATATGTCGGGAATGACTTCTCAGACCACTTTGACATTAAGTGGAGTGGCCGCAAT 1806
Db 1750 ACAAGATATGTCGGGAATGACTTCTCAGACCACTTTGACATTAAGTGGAGTGGCCGCAAT 1809
Qy 1807 AACGTAGAAATGATATTTATTTAATTGCCCCAAAGTGAACATATTCATTGAGGTGCA 1866
Db 1810 AACGTAGAAATGATATTTATTTAATGACCAAAAGGGAGCGTATTCATTGAGGTGACAG 1869
Qy 1867 GCATATATGTCGCGGTGGACCACAAACTTCTCGTGGCAATTGTGAAGCTTA 1920
Db 1870 GCTTATTAAGTACGATCGGTGGACCAAGACCTTCTCGTGGCAATTGTGAATTA 1923

RESULT 9
AR368118
LOCUS AR368118 .1
DEFINITION Sequence 7 from patent US 6376227.
ACCESSION AR368118
VERSION AR368118.1 GI:34601779
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kaseyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 7 23-APR-2002;
Kao Corporation; Tokyo;
JPK;

FEATURES
source
1..1923
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 76.9%; Score 1477.2; DB 6; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Db 430 TATTTGCGGATATACGAATAGATCCCAAGCTTTTCAAAAGGGGCAATCAGAGCTTGA 489
Qy 487 AAAGCTTGGCGCTTGATTCGAAGCAAAATTAAGAGTGCATTAAGAGCATCGAG 546
Db 490 AAAGAGTGGCGCTTGATTCGAAGCAAAATTAAGAGTGCATTAAGAGCATCGAA 549
Qy 547 GAAATGCGCTAGTACGTAGCAAGCAATGACGTCCTTATATTAACGGCAAAAGCTGAAT 606
Db 550 CAAATGCGCAATTTGCAATTAAGCAATGATGCTATATTAACGGCAAAAGCTGAAT 609
Qy 607 AAGGTAGTAATGATGTCGCGCAGAGTATTTGCAAGCGAGTGGCAAGAGCAAGCTAC 666
Db 610 AAGGTAGTAATGATGTCGCGCTGGAATTTGCAAGCGAGTGGCTCAGAGCAAGCTAC 669
Qy 667 GGTGTATGCAAAAGCGCAAGTTGTGCGAGTTGCCGATCTGGAATTGATACAGAGAA 726
Db 670 GCGTGTATGCAAAAGCAGATCGTAGCGGTGCGGATACAGGGCTTGATACAGGTCCG 729
Qy 727 AACGACGTTGATGATGAAGACCTTCGCGGTAAATTAACAGCACTATATGACATGGGT 786
Db 730 AATGACAGTTGATGATGAAGCTTCGCGGAAATTAATGCACTATATGACATTTGGGA 789
Qy 787 CGACGAATTAATGCAATGATGAACAGGTCAATGTTACCATGTGCGAGGTTCCGTAATTA 846
Db 790 CGGACGAATTAATGCAATGATGAATGATGTCATGTTACGATGTCGAGTGGCTCCGATTA 849
Qy 847 GGAATATGCGCAACGAATTAAGAAATGGCACTCAAGCGAATCTGGTTTTCATATCATC 906
Db 850 GGAATATGCGCTCAATTAAGAAATGGCGCTCAGCGAATCTGATCTTCAATATCATC 909
Qy 907 ATGATTAACAGTGTGGGCTTGGAGGCTTGGACCTTCCATATTCGCAAACTTATACGCCAA 966
Db 910 ATGATTAACAGTGTGGGCTTGGAGGCTTGGACCTTCCATATTCGCAAACTTATACGCCAA 969
Qy 967 GCATTCAGTGCAGTGCAGAAATTCATCAAACTCTGCGGGGAGCGGTAATGCGGCG 1026
Db 970 GCATTCAGTGCAGTGCAGAAATTCATCAAACTCTGCGGGAGCGAGTGAATGGGCT 1029
Qy 1027 TACAGCAAGATTCAGAAATGTGATGATCTATGTAGAAATTAATGATGATGATTTCT 1086
Db 1030 TACAGCAAGATTCAGAAATGTGATGATCTATGTAGCAAAATTAATGATGATGATTTCT 1089
Qy 1087 TTGCGGCTGGGAATGAAGAGCGGCAACGCGGTACCATTCAGTGCACCTGTACCGCTAA 1146
Db 1090 TTGCGTGGCGGAATGAAGAGCGGCAACGCGGTACCATTCAGTGCACCACTTAA 1149
Qy 1147 AACGCATTAACAGTGGCGCAACGAAACCTGCGCTCAAGGCTTCCGATTCATGACAT 1206
Db 1150 AATGCAATTAACAGTGGAGCTACGAAACCTGCGCTCAAGGCTTCCGATTCATGACAT 1209
Qy 1207 AATATTTAACAGTGTGCAAGTTCTTCCGCTGCGCCGCAAAAGATGGCGAATCAAG 1266
Db 1210 AATATCAACAGTGTGCAAGTTCTTCCGCTGCGCCGCAAAAGATGGAGATCAAA 1269
Qy 1267 CCGATGTCATGGCGCCGAGGACATCACTTTATGACGAAGTCTTCTTGCACCCGAT 1326
Db 1270 CCGATGTCATGGCGCCGAGGACATCACTATGACGAAGTCTTCTTGCACCCGAT 1329
Qy 1327 TCCCTCTTCTGGGCGAATCATGACAGCAATATGCTATATGAGGTAACGTCATGCA 1386
Db 1330 TCCCTCTTCTGGGCGAATCATGACAGTATGATATGATATGAGGTAACGTCATGCT 1389
Qy 1387 ACACGATTTGTTGCGGGAAATGTTCGACAGCTCCGCTGAGCAATTTGTGAAAATAGAGA 1446
Db 1390 ACACGATCGTTGCTGGAACGTTGCAAGCTTCGCTGAGCAATTTGTGAAAATAGAGG 1449
Qy 1447 ATCACTCTTAAGCTTCCCTATTTGAAGACAGCTTTGATTTGACAGTGTGCTGATTTGA 1506
Db 1450 ATCACTCAAAAGCCCTTCTATTTAAAGCGGCACTGATTTGCGGCTGACGTCGATCGGC 1509
Qy 1507 TTGGGTATTCGGAACGAAACCAAGAGTGGGCGAGTGAACCTGTGATTAATCGTTGAG 1566
Db 1510 CTTGGCTACCGAACGTTACCAAGATGGGAGCAAGTGAATTTGATTAATCCTGTAAC 1569

Oy	1567	GTTCGCTATGTGACGAAATCCAGTGCCTTATCAACTAGCCAAAGCAATATACCTTT	1626
Db	1570	GTTCGCTATGTGACGAGTCCAGTCTCTATCAACAGCAAAAGCAGATCTGTTT	1629
Oy	1627	ACTGGAACGGCGGGCGAAGCCATTGAAATCTCCCTGGTATGGTCCGATGCCCCGCAAGC	1688
Db	1630	ACTGCTACTGCGCGGCAAGCTTTGAAAAATCTCCCTGGTATGGTCTGATGCCCCGCGAGC	1688
Oy	1687	ACTACTGCTCTGTGTAACCCCTGGTCATGATTTGAAATTTGGTATATACAGACCAAAACGSA	1748
Db	1690	ACAACTGCTTCCGTAAACGCTTGTCATGATCTGGAACCTTGTCATTAACCGCTCCAAATGSC	1749
Oy	1747	ACAAATATATGTCGGAGATGACTCTTCAGACCACTTTGACAAATTACTGGAGATGGCCGCAAT	1806
Db	1750	ACAAAGTATGTAGAGAAATGACTTATCTTGCCATATCAATGATATATGGAGATGGCCGCAAT	1809
Oy	1807	AACGTAGAAATGATATTTATTAATTCGCCCCCAAGTGAACATATATCAATTTGAGGTGCAT	1866
Db	1810	AACGTAGAAATGATATTTATTAATTCACCAACAAACGGGAGCGTATATCAATTTGAGATGACG	1869
Oy	1867	GCATATATATGTCGCGGTGGACCAACAAACTCTCTGTGGCAATTTGTAATTA	1920
Db	1870	GCTATATATCGTACCGGTGGACCAACGAATCTCTGTGGCAATTTGTAATTA	1923

RESULT 10
AR562465

LOCUS	AR562465	1923 bp	DNA	linear	PAT 08-OCT-2004
DEFINITION	Sequence 5 from patent US 6759228.				
ACCESSION	AR562465				
VERSION	AR562465.1	GI:53976506			

SOURCE
ORGANIS

REFERENCE
AUTHORS
TITLE
JOURNAL

Unclassified.
1 (bases 1 to 1923)
Takaizawa, M., Okuda, K., Saeki, K., Kubota, H., Hitomi, J., Kageyama, Y.,
Shtkara, S., and Nomura, M.
Alkaline protease
Patent: US 675928-A 5 06-JUL-2004;
Kao Corporation; Tokyo,
JPN;

FEATURES

```

/organism="unknown"
/mol_type="genomic DNA"

```

ORIGIN

Query Match	76.9%	Score 1477.2	DB 6	Length 1923
Best Local Similarity	85.7%	Pred. No. 0		
Matches 1641	Conservative 0	Mismatches 273	Indels 0	Gaps 0
QY	7 AAGAAGAGGTTTATCTGTTTATCAGCTGACGAGATTCTGTGACGTGCATTA	66		
DB	10 AAGAAAAAGGTTTTTATCTGTTTATCAGCTGACGAGATTCTGTGACGTGCATTA	69		
QY	67 AACATCCCTCGGCTGATGATGCAAGACTTWTGATCTGATTTTAAAGAAATCAACA	126		
DB	70 AGTATTCATCTGCAGGCTGTCAGAGAAATTTGATCTGAAATTCAGAAATTCAGACA	129		
QY	127 ACAACCGATGTCAGTGTTTCTCCAAACAGCAGCAACACGCTGCGCTCATTTCTGGTG	186		
DB	130 ACAATGATGCTTAAAGTTTCTCCAAACAGGGGAGCATGGTGTGCTGCTTTTCTGGTG	189		
QY	187 GAGTCTGAAAATGTGAAAATTCTTAAAGATGCTTAAAGAAATTGAAACAGTACCGGCA	246		
DB	190 GAATCTGAAAATGTGAAATCCCAAAAGTTTGACAGAGAACTTGAACAGTCCCGGCA	249		
QY	247 AATATAAATCTCCATATTGTGCAATTCATAGGCCCATTTTAAAGAAACAAACAGAG	306		
DB	250 AATATAAATCTCCATATTGTGCAATTCATAGGCCCATTTTAAAGAAACAAACAGAG	309		

QY	307	TTAAGAGAACTGGAGGAAAGAATTTCTGACATACATCCCTGATATATGCTATATTTGTCGAG	366
Db	310	CTGGAAAAACAAGGGGCAAAAGATTTCTGACATACATCTGATATATGCTTACATTTGTCCAG	369
QY	367	TATGAGGGGGATGTTCAGTCAAAAGTCCGCTCCATTTGAAACAGTGGATTCAGTGGAGCCA	426
Db	370	TATGAGGGCGATGTTTAATCTCAGCAACAAGCACATTTAGCAACGTGGATTCGCTGGAGCCT	429
QY	427	TACTTGGCGAAATACAAATATGATTCGCCAGCTTTTCAAAAAGCGCATGCACGCTGGTG	486
Db	430	TATTTGCCGATATACAGATATGATTCGCCAGCTTTTCAAAAAGGGGCGATCAGAGCTTGT	489
QY	487	AAAGCGTTGGGGCTTTGATTCGAGACGAAACAATTAAGAAATGTCATTTAAAGGCATCGAG	546
Db	490	AAAGCAGAGGGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCAATTAAGAGGCATCGAA	549
QY	547	GAATTCGCTCAGTACGTGCAAGGACATGACGTCAATATATTAAGGCAAAAGCTGAAAT	606
Db	550	CAATTCGACATTTGCGCATTAAGCAATATGTGCTATATATTAAGGCAAAAGCTGAGAT	609
QY	607	AAGGTATGAATGATGTGGCCAGAGTATTTGTCAAAGCGATGTGGCACAGACAGCTTAC	666
Db	610	AAGGTATGAATGATGTGGCGGAGATTTGTCAAAGCGATGTGGCTCAGAGCAGCTAC	669
QY	667	GCTTTGTTAGACAAGAGCCAGATTTGTCCAGTTGCCATCTCGATTTGATATACGAAGA	726
Db	670	GGTGTATAGACAAGAGCAGATGTATCGATTCGGTTGCCATTAAGAGCTTGAATACAGTTCGC	729
QY	727	AACGACAGTTGATGATGATGAAGACCTTCGGGGTAAATTAACAGACCTATATGACCTGGGT	786
Db	730	AATGACAGTTGATGATGATGAAGACCTTCGGGGAAATTAATCTGATTTATATGACTTGGGA	789
QY	787	CGACAGATTAATGCGAATGATACGACCGGTATGGTACCATGTGTGACAGTTCCGTATTA	846
Db	790	CGACAGATTAATGCGAATGATACGACATGATGATGATGATGATGATGATGATGATGATGAT	849
QY	847	GGAATATGGCCGACAGATTAAGAAATGGACCTCAAGCAATCTGTGTTTTCATTCATC	906
Db	850	GGAATATGGCCCTCAGATTAATGAAGAAATGGCGCTCAGAGGAAATCTATGTTTCCATCTATC	909
QY	907	ATGATATGACAGTGTGGGCTTGGAGGCTGCTTCCATCTGCGAAACCTTATTCAGCCAA	966
Db	910	ATGATATGACAGTGTGGGCGATTTGGAGGACTTCTTCCGAAATCTTTCAGAACTTATTCAGCCAA	969
QY	967	GCATTCAGTGCAGGTTGCCAGAAATTCATACAAACTCTGGGGGGGACGCGGTGAAATGGGGCC	1026
Db	970	GCATTCAGTGTGGTGGCCAGAAATTCATACAAACTCTGGGGGACGAGCATGAAATGGGGCT	1029
QY	1027	TACACGACAGATTCAGAAATGTGATGATCTATGTAAAGAAAAATGATATGACGATTTCTT	1086
Db	1030	TACACACAGATTCAGAAATGTGATGATCTATGTGCGCAAAAATGATATGACGATTTCTT	1089
QY	1087	TTTCCGGCTGGGAATGAAGAAGCCGAAACGGCGGTACATCAGTGCACCTGTGACCGGCTTAA	1146
Db	1090	TTTCCGGCTGGGAATGAAGAAGCCGAAACGGCGGTACATCAGTGCACCGGCGCTTAA	1149
QY	1147	AACCCCATTAAGTGTGGCGCAACGAAAACTGTGGTCCAGCTTCGGTCTCTATATGACAGAT	1206
Db	1150	AATGCATTAAGTGTGGCGCAACGAAAACTGTGGTCCAGCTTCGGTCTCTATATGCGAC	1209
QY	1207	AATATTAACAAGTTTGACAGTTCCTTCCCGTGGCCCGACAAAGAATGGCGAATCAAG	1266
Db	1210	AATATTAACAAGTTTGACAGTTCCTTCCCGTGGCCCGACAAAGAATGGCGAATCAAA	1269
QY	1267	CCTGATGTCAATGGGGCCAGGGACATCACTTTATTCAGCAAGATCTTCTCTTGCACCGGAT	1326
Db	1270	CCGATGTCAATGGGACCGGGAAACGTTCACTATTCAGCAAGATCTTCTCTTGCACCGGAT	1329
QY	1327	TCCTTCCTCTGGGGGAATCATGACAGCAAAATGATCTATATGGGTGGAACGTCATGGCA	1386
Db	1330	TCCTTCCTCTGGGGGAATCATGACAGCAAAATGATCTATATGGGTGGAACGTCATGGCT	1389
QY	1387	ACACCGAATGTGTGGGGGAATGTTCACAGCTCCGTGACATTTTGTGAAAAATTAAGGA	1446

Db 1390 ACACGGATGTTGCTGGAAACGTGGCAACGCTTCGTGAGCATTTCGTAAACACGAGGC 1445

Qy 1447 ATCATCTCTTAAGCCTTCCTCTATTGAAAGCAGTTGATTCGACAGTCTGCTGATTTGGA 1506

Db 1450 ATCACACCAAAACCTTCCTCTATTAAAGGGGACGATGATTCGCGGTGCAGCTGACATCTGGC 1503

Qy 1507 TTGGGTTATCCGACGGAACCAAGATGGGGCCGAGGACCCCTGGATTAATTCGTTGAC 1566

Db 1510 CTGGCTACCCGACGGTBAACCAAGATGGGGACGATGACATTGGATTAATTCCTGTAC 1565

Qy 1567 GTTGCCTATGTGAACGAATCCAGTCCAGTCCCTATCACTAGCCAAAAGCAGATATACCTT 1628

Db 1570 GTTGCCTATGTGAACGATCCAGTCCAGTCTCTATCAACACGACAAAAGCAGATATACCTT 1628

Qy 1627 ACTGCAACGGCGGGCGAAGCATTGAAATCTCCCTGTATGTCGATGCCCTTCGACAGC 1686

Db 1630 ACTGCTACTGCGCGGACGCTTTGAAATCTCCCTGTATGTCGATGCCCTTCGACAGC 1683

Qy 1687 ACTACTGCTCTGTACCCCTGGTCGATATTTGATTTGGTCATTACGACACCAAGCGA 1746

Db 1690 ACAACTGCTTCGCTACGCTTGTCAATGATCTGAGACTTGTCAATACGCTCCCAATGGC 1749

Qy 1747 ACAAGATATGTCGGAGATGACTTCTCAGACCATTTGACAATTAACCTGGAGATGGCCGCAAT 1806

Db 1750 ACAAGATATGTAGGAATAGACTTCACTTCGCATACATGATTAATGGGATGGCCGCAAT 1803

Qy 1807 AACGTAGAAAATGTATTTATTAATTGCCCCCAAGGTGACATATATCAATTGAGGTGCA 1866

Db 1810 AACGTAGAAAATGTATTTATTAATGCAACAAGCCGGAAGTATACATTTGAGGTGACG 1869

Qy 1867 GCATATATATGTCGGGTGGAACAACAACCTTCCTGTTGGGCAATTTGTAACCTAA 1920

Db 1870 GCTTATATGTAACGGGTGGAACAACAACCTTCCTGTTGGGCAATTTGTAATTTAA 1923

RESULT 11				
AR562466				
LOCUS	AR562466	1923 bp	DNA	linear
DEFINITION	Sequence 7 from patent US 6759228.			
ACCESSION	AR562466			
VERSION	AR562466.1	GI:53976507		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1923)			
	Takata,M., Okuda,M., Saeki,K., Kubota,H., Hicomi,J., Kageyama,Y.,			
	Shikata,S. and Nomura,M.			
TITLE	Alkaline protease			
JOURNAL	Patent: US 6759228-A 7 06-JUL-2004;			
	Kao Corporation; Tokyo;			
	JPX;			

Query Match	76.9%	Score 1477.2	DB 6	Length 1923
Best Local Similarity	85.7%	Pred. No. 0		
Matches 1641; Conservative	0	Mismatches 273	Indels 0	Gaps 0

QY 7 AAGAAAGAGGTGTTTATCTGTTTATCAGGTGAGGAACTCTGTCGACTGTGGATTA 66
Db 10 AAGAAAAAGGTGTTTATCTGTTTATCAGGTGAGGAACTTGTGACTGTGGCTTA 69
QY 67 AACATCCCTGGCTGTGTGATCAAGACATTGTGATCTGTGATTTTAAAGAAATTCAAACA 126
Db 70 AGTATTCATCTGCAGGTGTGTCAAGAAATTTGTGATCTGTGATTTCAAGAAATTCAGACA 129
QY 127 ACAACGATTCAGTGTGTTTCTCCAAACAGGACAAACAGGTGCGGCTGCAATTTCTGGTGTG 186

Db	130	ACACGAGTACGTAAGAGTTTCTCCAGACAGGGGACAGCTGCGTCTGCTTTTCTGCGG	183
Qy	187	GAGTCTGAAAATGTGAAACTTCTTAAGAGATTGCTAAAGAACTTGAACAGTACCGCA	246
Db	190	GATCTGAAAAATGTGAACTCCCAAAGGTTTGCAGAGAAAGAACTTGAACAGTCCCGCA	249
Qy	247	AATTAATTAATCTCATATTTGTCCAAATTCGATGGCCCCATTTTAGAGAAACAAACGAAG	306
Db	250	AATTAATTAATCTCATATTTATTCAAATTCAGATGACCAATTTTAGAAGAAACAAACGACAG	309
Qy	307	CTAGAGACACATCGAGACAAAGATTCTCGACTATCCCTGATTAATGACATATTTGTCGAG	366
Db	310	CTGCAAAAAACAGGGGCAAGATTTCTCGACTATACCTGATTAATGCTTAATTTGTCGAG	369
Qy	367	TATGAGGGGGATGTTCACTGCAAAAGTCCGCTCATTTGAACAGTGGAAATCAGTGAAGCCA	426
Db	370	TATGAGGGGGATGTTAAGTCAGCAACAGCACATTTGAGCAGTGGAAATCCGTCGAGCCT	429
Qy	427	TACTTGGCGGAATACAAATATAGATCCCGAGCTTTTCACAAAGGGGCACTGACCTGGTG	486
Db	430	TATTTTCCGATTTACAGAAATAGATCCCGAGCTTTTCACAAAGGGGCACTCAGAGCTTGTGA	489
Qy	487	AAAGCGTTGGCGCTTGATACGAAGACAGCAATTAAGAAGTGCATTTAAGAGGCAATCGAG	546
Db	490	AAAGCAGTGGCGCTTGATACAGACAGAAATTAAGAAGTGCATTTAAGAGGCAATCGAA	549
Qy	547	GAATTCGCTCACTAGTACGTAGCAGACATGACGTCCATTAATATACGGCAAGCCTGATAT	606
Db	550	CAATTCGACATTCGCAATAGCAATGATGTGCTATATATTAACGGCAAGCCTGATAT	609
Qy	607	AAAGTGAATGATGATGTGGCCAGAGGTATTTGTCAAAAGCGAATGTGGCACAGCAGCTAC	666
Db	610	AAAGTGAATGAAAGATGTTGGCGCTGGAATTTGTCAAAAGCGAATGTGGCTCAGAGCAGCTAC	669
Qy	667	GGTTTGTATGACCAAGGGCCGAATTTGTGCAAGTTGGCGAATACGATGTTGGATACAGAGAA	726
Db	670	GGTTTGTATGACCAAGAGACAGATCTGATGGGTGGCGATACAGGGCTTGATACAGGTCCG	729
Qy	727	AACGACAGTTCGATCAGATGAAGCCTTCGCGGTGAATTAACAGCACTATATGCACTGGGT	786
Db	730	AATGACAGTTCGATCAGATGAAGCCTTCGCGGGAAATTAATCGCATATATATGCACTGGGA	789
Qy	787	CGAGCGAATTAATGCGAATGATACGAACGCTCATGGTACCATGTGGCAGGTTCCGATATTA	846
Db	790	CGAGCGAATTAATGCGAATGATACGAATGTCATGGTACGCAATGTGGCTCCGATATTA	849
Qy	847	GGAATATGGCGCAACGATTAAGAAATGGCACTTCAAGGGAATTCGGTTTTTCAATCCATC	906
Db	850	GGAATATGGCTCCACTTAATTAAGAAATGGCGCTCAGGGGAATCTAATCTTCCAAATCTATC	909
Qy	907	ATGATATACAGTGTGGGCTTGGAGGCTTGCGCTTCATCTGCAAACTTATTCAGCCAA	966
Db	910	ATGATATACCGGTGGGGGAACTTGGAGAACTTACTTGCATCTGCAAACTTATTCAGCCAA	969
Qy	967	GCATTCACTGACAGTGCAGGAATTCATACAACTTCCTGGGGGGCAGCGGTGAATGGGGCC	1026
Db	970	GCATTCACTGCTGGGTGCCAGAAATTCATCAAACTCTCTGGGGGACGAGTGAATGGGGCT	1029
Qy	1027	TACACGACAGATTCAGAAATGTGTGATGATATGTAAGAAAAATGATATATGACATTTCTT	1086
Db	1030	TACACAAACAGATTCAGAAATGTGTGATGATATGTGACCAAAATGATATATGACATTTCTT	1089
Qy	1087	TTGCGGCGTGGGAATGAAGAGCGGACGGAATCGGCGGTACCATCAGTGCACCTGATACGGCTAA	1146
Db	1090	TTGCGTGC CGGAATGAAGAGCCGACGCGGAACATCAAGTGCACCTGATACGGCTAA	1149
Qy	1147	AAACGCAATACAGTTCGCGCAACGAAAACTGCGTCCAGCTTCGTTCTTATGACAGAT	1206
Db	1150	AATGCAATACAGTTCGCGCAACGAAAACTCGCGCCAACTTTGGGTCTTATATGCGCAG	1209
Qy	1207	AATATTAACACGTTTGCACTGTTCTTCTCCGTGGCCGCAAAAAAGATGGCGCAATCAAG	1266
Db	1210	AATATCAACACATGTGGCACTGTTCTTCTTCACTGTGACGGAAGAAAGATGACGATTCAA	1269

QY 1257 CCTGATGTCATGCGCGGCAAGGACATATTTATGACGAAGTCTTCTCTGACCCCGAT 1326
DB 1270 CCGGATGTCATGCGCGGCAAGGACATATTTATGACGAAGTCTTCTCTGACCCCGAT 1329
QY 1327 TCCCTCTTGGGGGGAATGATGACGAAATATGCTATATGAGTGAACGTCATGCGA 1386
DB 1330 TCCCTCTTGGGGGGAATGATGACGAAATATGCTATATGAGTGAACGTCATGCGA 1389
QY 1387 ACACGATTTGTGCGGGGAATGTCACAGCTCCGTCGACATTTTGTGAAAAATAGAGA 1446
DB 1390 ACACGATTTGTGCGGGGAATGTCACAGCTCCGTCGACATTTTGTGAAAAATAGAGA 1449
QY 1447 ATCACTCTTACGCTTCCCTATTTGAAGACGTTGATGTCGCTGCTGATGTTGA 1506
DB 1450 ATCACTCTTACGCTTCCCTATTTGAAGACGTTGATGTCGCTGCTGATGTTGA 1509
QY 1507 TTGGGTTATCGGAACGGAACCAAGGATGGGGCGAGTGAACCTGGAATTAATCGTTGAC 1566
DB 1510 CTGGCTACCGCAACGCTTCTATTTAAAGCGCATGATGTCGCTGCTGATGTTGA 1569
QY 1567 GTTGGCTATGTCGAACGATTCAGTCCCTATTCAGTACGCAAAAGCGACATATACCTTT 1626
DB 1570 GTTGGCTATGTCGAACGATTCAGTCCCTATTCAGTACGCAAAAGCGACATATACCTTT 1629
QY 1627 ACTGCAACGGCGCGGCAAGCATTTGAAATCTCCCTGCTATGCTGCTGCTGCTGCAAGC 1686
DB 1630 ACTGCTACTGCGCGGCAAGCATTTGAAATCTCCCTGCTATGCTGCTGCTGCTGCAAGC 1689
QY 1687 ACTACTGCTTCTGTATACCTGCTGCTATGATGTTGATGTTGATGATGATGATGATGAT 1746
DB 1690 ACTACTGCTTCTGTATACCTGCTGCTATGATGTTGATGTTGATGATGATGATGATGAT 1749
QY 1747 ACAGATATGTCGGGAATGATCTTCTCAGACCAATTTGCAATTAATGCTGCTGCTGCTGCT 1806
DB 1750 ACAGATATGTCGGGAATGATCTTCTCAGACCAATTTGCAATTAATGCTGCTGCTGCTGCT 1809
QY 1807 AACGTAAGAAATGATATTTATTTATTTATGTCGCGCCGAAGTGAACATATTCATGAGTGCA 1866
DB 1810 AACGTAAGAAATGATATTTATTTATTTATGTCGCGCCGAAGTGAACATATTCATGAGTGCA 1869
QY 1867 GCATTAATATGTCGCGGTCGACCAAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 1870 GCATTAATATGTCGCGGTCGACCAAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923

RESULT 12
AR069954 3003 bp DNA linear PAT 18-FEB-2000
LOCUS AR069954
DEFINITION Sequence 41 from patent US 5891701.
ACCESSION AR069954
VERSION AR069954.1 GI:7220842
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3003)
AUTHORS Sloma, A. and Christianson, L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 41 06-APR-1999;
FEATURES
LOCATION/Qualifiers
SOURCE 1..3003
/organism="Unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 59.2%; Score 1136.8; DB 6; Length 3003;
Best Local Similarity 74.9%; Pred. No. 8e-300; Indels 3; Gaps 1;
Matches 1436; Conservative 0; Mismatches 477;
5 GAAGAAGAGAGGTTTATCTGTTTATCTGTTTATCTGCTGACGTCAGCGATTCGTGCACTGTGCAAT 64
DB 859 GATGGAAGAGGTTTATCTGTTTATCTGTTTATCTGCTGACGTCAGCTATTTGCTTCTGTTGCTT 918

QY 65 TAAACATCCCTCGGCTGCTGATGCAAGACCTTTGATCTGGAATTTTAAAGAAATTCAA 124
DB 919 TAAAGATGCTTCTTCTTAAATTTGAGGCGAACAATTTTGAATTTGACCTTTAAAGGAGATAGGA 978
QY 125 CAACACCGATGTCAGTGGTTTCTCAAAACAGCGCAAAACAGTGGCGGCTCATTTCTG 184
DB 979 CACTTACGCTAGAGAGGCTCCACCAAGCAAGAAAAACGGAAAGGCACTTTTCTTG 1038
QY 185 TGGAGTGAATAATGTGAACCTTCTTAAAGATTGCTTAAAGAACTTGAACAGTACCG 244
DB 1039 TAACTCTGAATAATGTGAATAATCCAAAGATATTCAAAAAGAACTGAAGATGATTCAG 1098
QY 245 CAATATTAATCTCATTTTCTCAATTTGATGTCGCCCATTTTGAAGAAACAAACGA 304
DB 1099 CGGATTAACACCTATATATCTTCAATTTGACGACCTATTTTGAAGAAACGCAATTC 1158
QY 305 AGTAGAGACAACTGAGCAAAAGATTCTGACTACATCCCTGATTAATGATATATTTGCG 364
DB 1159 AACTAGAGAAACGGAGCGGAAATTTCTGATTAATACAGATTTACGCTTATATTTGCG 1218
QY 365 AGTAGAGAGGATGTTCAATCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGAGC 424
DB 1219 AATATGATGGGATGTAAGGCGCTAATCAAGCAATTTGCGCATTTGGAATCGGTTGAC 1278
QY 425 CATACTTCCGAATATCAAAATTAATATCCCGCTTTTCAAAAGCGGCAATCGACGCTG 484
DB 1279 CATATTTTAACTTTTATTAATAATGACCGCAATTAATTTTCAAGAGAGCTTCTGAATTA 1338
QY 485 TGAAGCGTTGGCGCTTGAATGACAGAGCAACATTAAGAAAGTGCATTAAGAGGATCG 544
DB 1339 TGAAGAGTATGCTTTATTAATAAGAAAGAAATTAAGAAAGTATGTTAAGAGATTCG 1398
QY 545 AGGAATGCTCAGTACGTAAGCAAGATGATGCTCATTTATTTACGGCAAGCTGAT 604
DB 1399 AACAATGCGCAATATGACGCAAAATATGATGATTAATGATTAACCCCAAGGCTGAT 1458
QY 605 ATAAGTGAATGAATGATGTCGAGATTTGTCAAAGCGGATGTCAGAGCGCT 664
DB 1459 ACGAAGTTTAAATGACCTGCGCGCTGATGTAAGAGCAGCGTCGCAAAATTAATCT 1518
QY 665 ACGGTTTATGACCAAGGCGATTTGTCAGTTCGCGATTCGATTCGATTCAGGAA 724
DB 1519 TTGGCTTATATGACAGAGAGAGATTTGACAGATTTGCTGATATCTGCGCTTATACAGGA 1578
QY 725 GAAACGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
DB 1579 GAAATGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638
QY 785 GTGCGAGCAATTAATGCAATGATGCAAGCGTCAATGATGATGATGATGATGATGATGAT 844
DB 1639 GCAAGCAATTAATGCAATGATGCAAGCGTCAATGATGATGATGATGATGATGATGATGAT 1698
QY 845 TAGGAATGCGGCAACGAATTAAGAAATGCACTCAAGCAATCTGTTTTCATTCGA 904
DB 1699 TAGGAAT--GTAACAAATTAAGGAATGCAACGGAACCAATCTGTTTTCATTCGA 1755
QY 905 TCATGATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
DB 1756 TTAATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815
QY 965 AAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1024
DB 1816 AAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
QY 1025 CTAACAGCAGATTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1084
DB 1876 CTAATAGCAGATCTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
QY 1085 TTTTTCGCGCTGGAATGAAGCGCAAGCGGCTGATGATGATGATGATGATGATGATGATGAT 1144
DB 1936 TTTTTCGCGCGGAAATGAGGACAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1995

QY	845	TAGGAAATGGCGCAACGAATAAAGGAATGGCACTCAAGCGAATCGGTTTCAATCA	904
Db	1639	TAGGAAT---GCTCAAAATTAAGGATGGACCGCAACCAATCTTAGTCTTCAATCTA	1755
QY	905	TCATGGAATAGCAGTGGTGGCTTGGAGCTTGACCTTCCATCTGCAACCTTATTACGCC	964
Db	1756	TTATGGAATAGTGGTGGAGGGCTGGGAGGACTCTGCTAATCTACAAACATTAATTCAGTC	1811
QY	965	AAGCATTCAGTGCAGTGGCAGAAATTCATACAACTCCGCGGGGGCAGCGGTGAATGGG	1022
Db	1816	AAGCATATATAGCTGGAGCGAATAATTCATACAAATTCATGCGGGGCTCCAGTAACCGGTG	1877
QY	1025	CCTACACGACAGATTCACAGAAATGTGATGACTATGTAAAGAAAAATGATATACGATTC	1084
Db	1876	CCTATACGACAGACTCTCGAAATGTGATGATTAATGTGAAAAAATGATATGACGATTC	1933
QY	1085	TTTTGCGGCTGGGAATGAAGAGCGCGGATCCATCAGTGCACCTGTGACGCTTA	1144
Db	1936	TTTTTGGCGGCCGAATATAGGGAGCAGGTAGCGGTATCAATCAGTGCACAGGAACAGCA	1999
QY	1145	AAAAAGCCATACAGTGGCGCGAACCGAAACTGTGTCAGAGCTTGGTCTCTATGCGAC	1204
Db	1996	AAAAAGCCATATCAGTGGCGCGAACCGAAACTGTGTCAGAGCTTGGTCTCTATGCGG	2055
QY	1205	ATTAATATTAAACCAAGTTGACAGTCTCTTCCGTGGCCCGACAAAGATGGCGAATCA	1264
Db	2056	ATTAATATTAAACCAATGTGCTCAATCTCTTCAAGAGGTCTCTAGAGATGACGTAATTA	2111
QY	1265	AGCCGATAGTCATGGCGCCAGGGACATACATTTATCAGCAAGATCTTCTCTTGACCCG	1322
Db	2116	AGCCGAGGTATGGGACCAAGGTACGTATATCTCTGCTAGATCATCATTAAGCTCCAG	2177
QY	1335	ATTCCCTCCTCTGCGCGCATCATGACAGCAATATGCTATATGGGTGGAGCGTCAATGG	1384
Db	2176	ATTCCCTCATTTCTGGGCAACCATGATATGTAATATGCTCATCATGAGTGGTAACTTCTATGG	2233
QY	1385	CAACACCGATTTGTGGGGGAATGTTGCAAGCTCCGTGACATTTTGTGAAAAATATAG	1444
Db	2236	CTACTCCAAATGTATACAGGTATATGTTGCAATTAAGGAGGCAATTTGTGAAAAATATAG	2299
QY	1445	GAATCCTCTCTAAGCCTTCCATTTGAAGAAGCAGCTTGTATGACAGGTGCTGCTGATGTTG	1504
Db	2236	GGGTATCTCTAAGCCTTCCATTTGAAGCTGCTTAAATTCAGAGTGTGCGGATGTTG	2355
QY	1505	GATGGGGTTATCGAACCGGAAACCAAGATGGGGCGGACGTGACCTTGATTAATCGTTGA	1564
Db	2356	GACTTGGCTTTCCAATGGTATACCAAGATGGGGGAAGTATCGTTAAGTAAATCCCTPA	2411
QY	1565	ACGTTGCCATATGTGAACGAATCCATGCCCCCTATGCACTTACCAAAAAGGACATATACCT	1622
Db	2416	ATGTGCAATTTGTGAATGAAGAGCCCTTATACAAAGTCAAAAAGACATATATTCGT	2477
QY	1625	TTACTGCAACGGCGGGCAAGCCATTTGAANAATCTCCCTGATATGTCGATGCCCCCTGCA	1684
Db	2476	TTAGGGCTCAACGTGTGTAACCTCTTAAAAATATCACTTTGTGTCAAGTGCACAGGTA	2533
QY	1685	GCACTACTGCTCTGTGAACCTGTGTCATGATTTGGATTTGGTCATATACGACCAAAACG	1744
Db	2536	GCAACACCGGCATCACTTAATTTAGGAAGATTTTGAATTAATATCACTGCACCAAAATG	2599
QY	1745	GAACGAAGATATGCGGAATGACTCTCAGCACCAATTTGACAAATACATTAACATTAAGTGC	1804
Db	2596	GAACTAAATACGTGGAAATGACTTTACAGCACCGTATATTAACAATTTGGATTTGGCAGAA	2655
QY	1805	ATTAACGTAGAAAATGATTTATTAATTCGCCCCCAAGTGAACATATACATTAAGTGC	1864
Db	2656	ACAAAGTGGAAAATGTGTTATCATATGCTCTCAAGCGGAACGTATACAGTCGAAGTGC	2711
QY	1865	AAGCATATATCTGCCGTTGGACCACAAAATCTCTCGTTGCAATTTGGAACCTTA	1920
Db	2716	AGGCTTACAAATATACAGTAATCCGCAAACTTTTCTTATACGANTGTACATTTAA	2771

[illegible]

QY 1096 GGGAAATGAAGGCGGAAACCGGCTACATGACGTGCTAGTCCGCTAATAAAGCCGATA 1155
 DB 481 GGGAAATGAAGGCGGAAACCGGCTACATGACGTGCTAGTCCGCTAATAAAGCCGATA 540
 QY 1156 ACAAGTCGGGCGCAACCGAAACCTGCGTCAAGCTTGGCTTCTATGCAATATATTAAC 1215
 DB 541 ACAAGTCGGGCGCAACCGAAACCTGCGTCAAGCTTGGCTTCTATGCAATATATTAAC 600
 QY 1216 CACGTCGCAAGTCTCTCCGCGGCGGCAAAAGATGGCGCAATCAAGCTGATGTC 1275
 DB 601 CACGTCGCAAGTCTCTCCGCGGCGGCAAAAGATGGCGCAATCAAGCTGATGTC 660
 QY 1276 ATGGCGCGGCGGCAATATATTAAGAGATCTCTCTGCAACCGATTCCTCTC 1335
 DB 661 ATGGCGCGGCGGCAATATATTAAGAGATCTCTCTGCAACCGATTCCTCTC 720
 QY 1336 TGGGCGAATCATGACAGCAAAATATGCTATATGGGTGAACGTCCATGCAACCGAT 1395
 DB 721 TGGGCGAATCATGACAGCAAAATATGCTATATGGGTGAACGTCCATGCAACCGAT 780
 QY 1396 GTTGGCGGGAATGTTGCAACGCTCCGTGACATTTTGTGAAAATATGAGAACTACTCT 1455
 DB 781 GTTGGCGGGAATGTTGCAACGCTCCGTGACATTTTGTGAAAATATGAGAACTACTCT 840
 QY 1456 AAGCCTCCCTATGAAAGCAGCTTGTATGACAGGCTGCTGATGTTGGATTGGTTAT 1515
 DB 841 AAGCCTCCCTATGAAAGCAGCTTGTATGACAGGCTGCTGATGTTGGATTGGTTAT 900
 QY 1516 CGGAACGGAACCAAGATGAGGCGGAGTGAACCTGATTAATTCGTTGAACGTTGCTAT 1575
 DB 901 CGGAACGGAACCAAGATGAGGCGGAGTGAACCTGATTAATTCGTTGAACGTTGCTAT 960
 QY 1576 GTGAACGAATCCAGTGCCTTATCACTAGCAAAAGCGACATATACCTTTATCTGCAAG 1635
 DB 961 GTGAACGAATCCAGTGCCTTATCACTAGCAAAAGCGACATATACCTTTATCTGCAAG 1020
 QY 1636 GCGGCGAAGCCATGAAATCTCCGTGATAGTGGGATGCGCCGCAAGCACTATGCT 1695
 DB 1021 GCGGCGAAGCCATGAAATCTCCGTGATAGTGGGATGCGCCGCAAGCACTATGCT 1080
 QY 1696 TCTGTAACCTGCTGATATGATTTGATTTGCTATTAAGCAACCAAGCAAGATAT 1755
 DB 1081 TCTGTAACCTGCTGATATGATTTGATTTGCTATTAAGCAACCAAGCAAGATAT 1140
 QY 1756 GTGGGAATGACTTCTGACGACATTTGACATTAACCTGGATGCGCGCAATACGTAGA 1815
 DB 1141 GTGGGAATGACTTCTGACGACATTTGACATTAACCTGGATGCGCGCAATACGTAGA 1200
 QY 1816 AATGATTTATTAATTTGCGCCCAAGTGAACATATACATTTGAGGTGAACATTAAT 1875
 DB 1201 AATGATTTATTAATTTGCGCCCAAGTGAACATATACATTTGAGGTGAACATTAAT 1260
 QY 1876 GTGCGGTTGACACAAACTTCTGTTGGCAATTTGCACTAA 1920
 DB 1261 GTGCGGTTGACACAAACTTCTGTTGGCAATTTGCACTAA 1305

RESULT 15
 E03808 2218 bp DNA linear PAT 29-SEP-1997
 LOCUS E03808 DNA encoding alkaline protease Ya.
 DEFINITION E03808.1 GI:2172022
 ACCESSION E03808.1 GI:2172022
 VERSION JP 1992197182-A/1.
 KEYWORDS Bacillus sp.
 SOURCE Bacillus sp.
 ORGANISM Bacillus sp.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1 (bases 1 to 2218)
 AUTHORS Tobe, S., Odera, M. and Arai, Y.
 TITLE DNA CODING ALKALINE PROTEASE YA ENZYME AND PRODUCTION OF ALKALINE
 PROTEASE YA USING THE DNA
 JOURNAL Patent: JP 1992197182-A 1 16-JUL-1992;
 LION CORP

COMMENT OS Bacillus sp.
 PN JP 1992197182-A/1
 PD 16-JUL-1992
 PF 28-NOV-1990 JP 1990327110
 PI TOBE SEICHI, ODERA MOTOYASU, ARAI YOSHIO
 PC C12N15/57, C11D3/366, C12N9/54, (C12N15/57, C12R1/07), (C12N9/54,
 C12R1/07);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: strain=Y strain;
 CC *source: clone=pUB8A;
 FH key location/qualifiers
 FH CDS 218..2125
 FT sig_peptide 218..823
 FT mat_peptide 824..2122
 FT /product= Precursor of alkaline protease Ya'.
 FEATURES
 source location/qualifiers
 1..2218
 /organism="Bacillus sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:1409"
 ORIGIN
 Query Match 48.4%; Score 930; DB 6; Length 2218;
 Best Local Similarity 70.5%; Pred. No. 3,76-243;
 Matches 186; Conservative 0; Mismatches 530; Indels 9; Gaps 3;
 QY 96 TTTGATCTGATTTTAAAGATTCAAACAAACCGATGTCAGTGTCTCCAAACA 155
 DB 310 TTTGATCTGATTTTAAAGATTTAAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAG 366
 QY 156 GCGAACAAGGTGGGCTGCACTTCTGTGAGTCTGAAAATGGAACCTCTAAAG 215
 DB 367 GATTAAGTGGGTGGGCACTCTTCTGATGATATGGAATATTAATATTCCTAAAG 426
 QY 216 ATTGCTAAAGAACTTGAAGACGTACCGGCAATATTAATCTCATATTTGCCAATTCA 275
 DB 427 TATTCAAAAGAAAGTGAAGGAGTACAGAAAGATACGAACCTCATCTGATCAATTTAC 486
 QY 276 TGGCCCAATTTGAAGAAACAAACAGAGCTAGAGCAACTGAGCAAGATTCTGCA 335
 DB 487 TGAACCAATTTCAAGAGAAAGCGAAAGATTAAGCTCTAGAGATCATCTCTAGA 546
 QY 336 CTACATCCCTGATTAATGATATATTTGAGATAGAGGGGATGTTCACTCAAAAGTCCG 395
 DB 547 TTATGTTCCAGATTAATGCTTTTATTTGTCAGAT--AGTGTGCTACAAAATATTAAG 603
 QY 396 CTCCATGAACAGTGAATCAGTGAAGCCTACTTGCAGAAATCAAAATATGATCCCA 455
 DB 604 TACTTAACTCTGTGGAAGAGTACACATTTTACATTAATTAATATGATCTGCA 663
 QY 456 GCTTTTCAAAAGGCGATGACGCTGTGAAACCGTTGGCGCTTGAATGACAGACAA 515
 DB 664 GCTTTTCAAAAGGCGCTTCCAGCTTGTTCAGCGGCTTATTTAAATACAAACACGA 723
 QY 516 CAATTAAGAAGTGCATTAAGAGCATGAGAAATGCTCAGTACGTAGCAAGCAATGA 575
 DB 724 AATTAAGAAGTGAATTAATTCGCTTGAATGAGATGCTTCAATATGCTGCAAAATGA 783
 QY 576 GGTCCATATATTAAGCGAAAGCTGATATTAAGGTGATGATGATGTCGACAGATAT 635
 DB 784 TGTGTTTATATATACCAAAAGCCGAGTATAGAGTATGATATGATGACAGAGGAT 843
 QY 636 TGTCAAGCGAGTGTGACAGAGCAGTACGTTTGTATGACAAAGCCAGATTGTGC 695
 DB 844 AGTAAGAAGCTAGTTGACAAACAAATTAAGATTAATGAGCAAGGTCACTAGTTGC 903
 QY 696 AGTTGCCATATCTGATTTGATACAGAAAGCAAGTTCGATGATGAAGCTTCCG 755

Db 904 AGTAGCGGACACAGGCTTAGATACAGTCTGTAACGATAGTTCTATGATGAAGCATTCGG 963
 Qy 756 CGGTAAATAAACAAGCATATATGCACTGGTTCGACGAATAATGCGAATGATACGACGG 815
 Db 964 CGGGAATAATCACAGCTCTTATGCGGTAGGAAGAACTAATATGAGAGATCCGAATGG 1023
 Qy 816 TCATAGTACCAATGCGAGGTTGGGTATTTAGGAATGGCGCAAGATTAAGAAATGGC 875
 Db 1024 GCATGGCACACATGTAGACAGTTCGTACTTGGTAAT--GCTTTAATAATAAAGGAATGGC 1080
 Qy 876 ACCTCAAGCGAATCTGTTTTTCAATCCATCATGATAGCAGTGGGCTTGAAGGCTT 935
 Db 1081 TCCGGAAGCTAATCTAGTCTTCCATCTAATATGATAGCAGCGAGATTAAGTGGCTT 1140
 Qy 936 GCCTTCCAATCTGCAACCTTATTCAGCCAGACATTCAGTGCAGGTGCCAATTCATAC 995
 Db 1141 ACCATCGAATTAATAGCTTATTTAGTCAGACCTTGAAATGCTGGAGCAAGAAATTCATAC 1200
 Qy 996 AAACCTCGGGGGGCGAGGGTGAATGGGGCTTACACAGATTCAGAAATGTGATGA 1055
 Db 1201 TAACTGTTGGGAGCCCGCAATTAATGAGCGTACCTGTACCTGAGCAAGTGAATGA 1260
 Qy 1056 CTATGTAAAGAAAAATGATATGATGACGATTCCTTTGCGGCTGGGAAATGAAGCCGACGG 1115
 Db 1261 GTATGTTGAAATATATGATATGACCGTACTTTTGAGCTGGTAAATGAAGTCTTAATTC 1320
 Qy 1116 CGGTACCATCAGTGCACTGTGAGGCTTAAAGCCATTAACATGCGGCGCAACCGAAA 1175
 Db 1321 AGGAACAATTAATGTCTCCAGGTACAGCGAAAAATCTATTACGTCGCGCAACGAAAA 1380
 Qy 1176 COTGGTCCAGGCTCGGTTCTATGCAATTAATTTAACGACGTTGCAAGTTCCTTC 1235
 Db 1381 CTATGCCCAAGCTTCGGTTGATAGCAGATTAACCAAAATCAATTTGCAAAATTTTCATC 1440
 Qy 1236 CCGTGGCCCGACAAAGATGGCGAATCAAGCCTGATGTCATGCGCCGAGGACATACAT 1295
 Db 1441 GAGAGAGACTACGAGGAGATGAGCAATTAAGCCTGACGTTAACGCTCTGGAAATTTAT 1500
 Qy 1296 TTTATCAGCAAGATCTTCTCTTGCAACCGATTCCTCTTCTGGGCGAATCATGACAGCA 1355
 Db 1501 TTTATCAGCAGTTCCTCTTGAAGCTCCAGACTCTTCGTTTGGCGAATTAATACAGTAA 1560
 Qy 1356 ATATGCTATATAGGTTGGAACGTTCATGCAACCGATTCGTTGGCGGGAATGTTGACA 1415
 Db 1561 ATACGCGTATATGCGCGGTAACCTCCATGCGACACCTATTTGTCAGGGAATGTCGCGCA 1620
 Qy 1416 GCTCCGTAGCAATTTTGTGAAAAATAGAGAACTACTCTTAAGCCTTCCTTAATTGAAAGC 1475
 Db 1621 ATTAAGTAGGCAATTTTATTAATAAATAGAGGTATTACTCTTAAGCCTTCCTTAATAAAGC 1680
 Qy 1476 AGCTTTGATGAGGTGCTGCTGATGTTGGATTTCGGAACGGAACCAAGGATG 1535
 Db 1681 TGCACCTTATCGTGTGCTACTGATGTTGGTTAGGATATCCTAGTGTGACCAAGGCTG 1740
 Qy 1536 GGGCCGAGTGAACCTGATTAATCGTTGAAAGTTGCTATGTGAAGCAATCCAGTCCCT 1595
 Db 1741 GGGGCGGTACTCTAGATTAATCGTTAAATGTAGCGTATGTCAATGAAGCACTGCATT 1800
 Qy 1596 ATCAATGAGCCAAAAAGCGACATTAACCTTTATCTGCAACGCGGCGCAAGCCATGAAAT 1655
 Db 1801 AGCCACAGGACAAAAAGCAACGTATTCGTTCCAGCACAAAGGGGTAAACCTTTAAAAAT 1860
 Qy 1656 CTCCCTGTATGTGCGATGCCCCCTGCAAGCATTAAGTCTGTAACCTGTGCAATGA 1715
 Db 1861 CTCGTTAGTATGAGACAGATGCTCCGGAAGTACAACTGCATCTTATACCTAGTTAATGA 1920
 Qy 1716 TTTGATTTGTCATTAACGACCAACGGAACAAGATATGTGGGAATGACTCTCAGC 1775
 Db 1921 TTTAGATCTAGTTATTACTGCTCCGAATGACAAAATATGTAGGAATGATTTTAGTTA 1980
 Qy 1776 ACGATTTGACATTAACGAGATGCGCGCATTAACGTAGAAAAATGTAATTTAATTCGCC 1835
 Db 1981 TCCCTATGATATACTGGGATGATCGCAACATGTGGAACGTATTTAATAACGCTCC 2040

Qy 1836 CCAAGTGAACATATACCATTTGAGTGCAGACATATATATGTGCCGTTGACCAACAAA 1895
 Db 2041 GCAATCTGGAAAGTATATTAATTTGAGTTCAAGCGGTATATATGATCACTGTGCCACAGCG 2100
 Qy 1896 CTTCTGTTGGCAATTTGAACTAA 1920
 Db 2101 TTTCTCACTAGCTATGTAATTA 2125

Search completed: April 8, 2006, 06:48:23
 Job time : 9412.15 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 8, 2006, 00:55:07 ; Search time 2979.67 Seconds
(without alignments) 4294.503 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atgagagaagaagaaggctgtc.....cgttgcaatctgtaactaa 1920

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Labeling first 45 summaries

Database : N_Geneseq.21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002as:*
6: geneseqn2002bs:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2003ds:*
11: geneseqn2004as:*
12: geneseqn2004bs:*
13: geneseqn2004cs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1920	100.0	1920	2	AAx37277 Bacillus
2	1477.2	76.9	1923	2	AAx37278 Bacillus
3	1477.2	76.9	1923	2	AAx37279 Bacillus
4	1477.2	76.9	1923	12	ADm40772 DNA encod
5	1477.2	76.9	1923	13	ADm14428 Bacillus
6	1477.2	76.9	1923	13	ADT49605 Bacillus
7	1136.8	59.2	3003	2	AAV82382 Bacillus
8	1002.6	52.2	1305	11	ADY33779 Bacillus
9	1002.6	52.2	1305	11	ADL25803 Bacillus
10	747.6	38.9	1299	2	AAQ27516 Alkali-pr
11	48.2	2.5	1332	12	ADm99124 Bacterial
12	47.4	2.5	2000	8	ADA71938 Rice gene
13	47.4	2.5	2000	8	ADA71938 Rice gene
14	40.6	2.1	2000	11	ACL37108 Rice gene
15	40.6	2.1	522	4	AA55654 DNA encod
16	40.6	2.1	522	4	AA55668 DNA encod
17	40.6	2.1	1446	14	ADV16383 E. faecal
18	40.6	2.1	1494	10	ADH84700 Enterococ
19	40.6	2.1	8911	2	AAx13588 Enterococ

20	40	2.1	8911	6	AB599383	AB599383 Enterococ
21	39.6	2.1	2382	10	ADJ56506	ADJ56506 Human CDN
22	39.6	2.1	2947	11	ADP65683	ADP65683 Human trl
23	39.6	2.1	2947	11	ADP65766	ADP65766 Human trl
24	39.6	2.1	4153	12	ADQ22355	ADQ22355 Human sof
25	39.6	2.0	522	4	AA556648	AA556648 DNA encod
26	38.8	2.0	2000	11	ACL35887	ACL35887 Rice gene
27	38.6	2.0	1992	8	ACA47875	ACA47875 Prokaryot
28	38.4	2.0	522	4	AA556717	AA556717 DNA encod
29	38.4	2.0	522	4	AA556718	AA556718 DNA encod
30	38.4	2.0	522	4	AA556641	AA556641 DNA encod
31	38.4	2.0	522	4	AA556681	AA556681 DNA encod
32	38.4	2.0	522	4	AA556710	AA556710 DNA encod
33	38.4	2.0	522	4	AA556721	AA556721 DNA encod
34	38.4	2.0	522	4	AA556723	AA556723 DNA encod
35	38.4	2.0	522	4	AA556658	AA556658 DNA encod
36	38.4	2.0	522	4	AA556720	AA556720 DNA encod
37	38.4	2.0	522	4	AA556678	AA556678 DNA encod
38	38.4	2.0	522	4	AA556635	AA556635 DNA encod
39	38.4	2.0	522	4	AA556664	AA556664 DNA encod
40	38.4	2.0	522	4	AA556677	AA556677 DNA encod
41	38.4	2.0	522	4	AA556716	AA556716 DNA encod
42	38.4	2.0	522	4	AA556645	AA556645 DNA encod
43	38.4	2.0	522	4	AA556674	AA556674 DNA encod
44	38.4	2.0	522	4	AA556659	AA556659 DNA encod
45	38.4	2.0	522	4	AA556728	AA556728 DNA encod

ALIGNMENTS

RESULT 1
AAx37277
ID AAx37277 strand; DNA; 1920 BP.
XX
XX
AC AAx37277;
XX
XX
DT 20-MAR-2003 (revaled)
DT 21-JUL-1999 (first entry)
XX
XX
DB Bacillus alkaline protease encoding DNA.
XX
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidising agent; ss.
XX
XX
OS Bacillus sp.
XX
XX
PN MO9918218-AL.
XX
XX
PD 15-APR-1999.
XX
XX
PF 07-OCT-1998; 98WO-JP004528.
XX
XX
PR 07-OCT-1997; 97JP-00274570.
XX
XX
PI (KAOS) KAO CORP.
XX
XX
PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX Shikata S, Nomura M,
XX
XX
DR WPI, 1999-287736/27.
XX
XX
DR P-PSDB; AAx17087, AAx17089.
XX
XX
PS Alkali protease from Bacillus used in washing powders.
XX
XX
PS Disclosure; Page 53-58; 71pp; Japanese.
XX
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkali protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is

CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX

60 Sequence 1920 BP; 576 A; 411 C; 473 G; 460 T; 0 U; 0 Other;

Query Match 100.0%; Score 1920; DB 2; Length 1920;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGAAAAGAGAGGTGTTTATCTGTTTATCAGCTGCAGCGATCTGTGACTGTT 60
DB 1 ATGAGAAAAGAGAGGTGTTTATCTGTTTATCAGCTGCAGCGATCTGTGACTGTT 60
QY 61 GCATTAAACAATCCCTCGGCTGGTATGCAAGGACTTTTATCTGGATTTTAAAGAAAT 120
DB 61 GCATTAAACAATCCCTCGGCTGGTATGCAAGGACTTTTATCTGGATTTTAAAGAAAT 120
QY 121 CAACAACAACCGATGTCAGTGTCTTCCAAACAGCGACAAACAGGTGCGCTGCATT 180
DB 121 CAACAACAACCGATGTCAGTGTCTTCCAAACAGCGACAAACAGGTGCGCTGCATT 180
QY 181 CTGGTGGAGCTGGAATAATGTGAACCTTCTTAAAGATTCCTTAAAGAACTTGAACAGTA 240
DB 181 CTGGTGGAGCTGGAATAATGTGAACCTTCTTAAAGATTCCTTAAAGAACTTGAACAGTA 240
QY 241 CCGGCAATTAATACTCCATTTGTCCAAATTCATAGGCCCCATTTTGAAGAAACAAA 300
DB 241 CCGGCAATTAATACTCCATTTGTCCAAATTCATAGGCCCCATTTTGAAGAAACAAA 300
QY 301 CAGAACTAGAGACAATGAGCAAAAGATTCGACCTACATCCGATTAATGCAATATTT 360
DB 301 CAGAACTAGAGACAATGAGCAAAAGATTCGACCTACATCCGATTAATGCAATATTT 360
QY 361 GTCCGATGAGGGGGAGTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCATG 420
DB 361 GTCCGATGAGGGGGAGTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCATG 420
QY 421 GAGCCATCTTCCGAAATTAACAATAGATCCCAAGCTTTTCACAAAAGCGCATCGAG 480
DB 421 GAGCCATCTTCCGAAATTAACAATAGATCCCAAGCTTTTCACAAAAGCGCATCGAG 480
QY 481 CTGGTGAAGCGTGGCGCTTGATTCGAAGCAACAAATAAAGATGCAATTAAGAGGC 540
DB 481 CTGGTGAAGCGTGGCGCTTGATTCGAAGCAACAAATAAAGATGCAATTAAGAGGC 540
QY 541 ATCCGAGAAATCGCTCAGTAGCTAGCAAGCAATGACCTTATTAATTAACGCAAAAGCT 600
DB 541 ATCCGAGAAATCGCTCAGTAGCTAGCAAGCAATGACCTTATTAATTAACGCAAAAGCT 600
QY 601 GAATATAAGTATGATGATGTGGCCAGAGGTATTTCAAGCGCATGTGGCAGAGC 660
DB 601 GAATATAAGTATGATGATGTGGCCAGAGGTATTTCAAGCGCATGTGGCAGAGC 660
QY 661 AGCTACGTTTGTATGACAAAGCGCATGTCGAGTTGCCGATCTGAATTGGATACA 720
DB 661 AGCTACGTTTGTATGACAAAGCGCATGTCGAGTTGCCGATCTGAATTGGATACA 720
QY 721 GGAGAAACGACAGTGTGATGATGACAGCTTCCGCGTAAATTAACGCACTATATGCA 780
DB 721 GGAGAAACGACAGTGTGATGATGACAGCTTCCGCGTAAATTAACGCACTATATGCA 780
QY 781 CTGGGTGCGAGCAATATGCAATGATGCAACGCTCATGTTACCATGTGGCAGGTTCC 840
DB 781 CTGGGTGCGAGCAATATGCAATGATGCAACGCTCATGTTACCATGTGGCAGGTTCC 840
QY 841 GTATTAGAAATGGCGCAAGATTAAGGAATGGCACTTCAAGCGAATCTGCTTTTCA 900
DB 841 GTATTAGAAATGGCGCAAGATTAAGGAATGGCACTTCAAGCGAATCTGCTTTTCA 900
```

```
DB 841 GTATTAGAAATGGCGCAAGATTAAGGAATGGCACTTCAAGCGAATCTGCTTTTCA 900
QY 901 TCATCATGATATGAGATGAGTGGCTTGGAGGCTTGCCTTCAATCTGCAAACTTATTC 960
DB 901 TCATCATGATATGAGATGAGTGGCTTGGAGGCTTGCCTTCAATCTGCAAACTTATTC 960
QY 961 AGCCAAAGCATTCAGTGCAGGAGCCAGAAATTCATCAAACTCTGGGGGGGAGCGGTGAT 1020
DB 961 AGCCAAAGCATTCAGTGCAGGAGCCAGAAATTCATCAAACTCTGGGGGGGAGCGGTGAT 1020
QY 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATCTATGTAAGAAATATATATGACG 1080
DB 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATCTATGTAAGAAATATATATGACG 1080
QY 1081 ATTCTTTTCCGGCTGGGAATGAAAGCCGAACGGCGGTACATCATGTCACCTGTGACG 1140
DB 1081 ATTCTTTTCCGGCTGGGAATGAAAGCCGAACGGCGGTACATCATGTCACCTGTGACG 1140
QY 1141 GCTAAAAACGCGCATPACAGTGGGGCAACGAAACCTGGGCTGCAAGCTTCCGTAT 1200
DB 1141 GCTAAAAACGCGCATPACAGTGGGGCAACGAAACCTGGGCTGCAAGCTTCCGTAT 1200
QY 1201 GCAATATATATTAACAAGTTCACAGTTCCTTCCGCTGCCCCGACAAAGATGGGCGA 1260
DB 1201 GCAATATATATTAACAAGTTCACAGTTCCTTCCGCTGCCCCGACAAAGATGGGCGA 1260
QY 1261 ATCAAGCTGATGTCATGAGGCGCCAGGACATACATTTTATGACAAAGTCTTCTTGA 1320
DB 1261 ATCAAGCTGATGTCATGAGGCGCCAGGACATACATTTTATGACAAAGTCTTCTTGA 1320
QY 1321 CCCGATTCCTCTCTGCGGCGAATCATGACAGCAAAATATGCTTATGAGTGAACGTCC 1380
DB 1321 CCCGATTCCTCTCTGCGGCGAATCATGACAGCAAAATATGCTTATGAGTGAACGTCC 1380
QY 1381 ATGSCAACACCGAATGTTGTCGGGGAAATGTGACAGCTCGTGAGCATTTTGTGAAAAT 1440
DB 1381 ATGSCAACACCGAATGTTGTCGGGGAAATGTGACAGCTCGTGAGCATTTTGTGAAAAT 1440
QY 1441 AGAGAAATCACTCCTTAAGCCTTCCCTATTGAAAGCACTTTGATGAGGTGCTGAT 1500
DB 1441 AGAGAAATCACTCCTTAAGCCTTCCCTATTGAAAGCACTTTGATGAGGTGCTGAT 1500
QY 1501 GTTGATTTGGGTATCCGAAACGAAACCAAGATGGGGCGAGTGAACCTGGAATATTCG 1560
DB 1501 GTTGATTTGGGTATCCGAAACGAAACCAAGATGGGGCGAGTGAACCTGGAATATTCG 1560
QY 1561 TTGAACGTTGCTATGTGAACGAATCCAGTGCCCTATCAATGACCAAAAAGCGACATAT 1620
DB 1561 TTGAACGTTGCTATGTGAACGAATCCAGTGCCCTATCAATGACCAAAAAGCGACATAT 1620
QY 1621 ACCCTTACTGCAACGGGGCGAAGCCATTGAAAATCTCCCTGTGATGTGGATGCCCT 1680
DB 1621 ACCCTTACTGCAACGGGGCGAAGCCATTGAAAATCTCCCTGTGATGTGGATGCCCT 1680
QY 1681 GCAAGCACTACTGCTCTGTAAACCTGTGATGATTTGATTTGCAATTAACGACCA 1740
DB 1681 GCAAGCACTACTGCTCTGTAAACCTGTGATGATTTGATTTGCAATTAACGACCA 1740
QY 1741 AACGAAACAAAGATATGTGCGAATGACTTCTGACACATTTGAACAATTAACGAGTGC 1800
DB 1741 AACGAAACAAAGATATGTGCGAATGACTTCTGACACATTTGAACAATTAACGAGTGC 1800
QY 1801 CGCAATTAAGTGAAGAAATGTATTAATTAATTCGCCCAAGGTGAACATATACATTTGAG 1860
DB 1801 CGCAATTAAGTGAAGAAATGTATTAATTAATTCGCCCAAGGTGAACATATACATTTGAG 1860
QY 1861 GTGCAACATATATATGTGCGGTTGACCAAAACTTCTGTTGCAATTTGCAATTA 1920
DB 1861 GTGCAACATATATATGTGCGGTTGACCAAAACTTCTGTTGCAATTTGCAATTA 1920
```

RESULT 2
AAK37278

ID AAX37278 standard; DNA; 1923 BP.
XX
AC AAX37278: .
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.
XX
KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent; ss.
XX
OS Bacillus sp.
XX
PN MO918218-A1.
XX
PD 15-APR-1999.
XX
PF 07-OCT-1998; 98WO-JP04528.
XX
PR 07-OCT-1997; 97JP-00274570.
XX
PA (KAOS) KAO CORP.
XX
PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
PI
DR WPI; 1999-287736/27.
DR P-PSDB; AAY17088, AAY17090.
XX
XX
XX Alkali protease from Bacillus used in washing powders.
PS Disclosure; Page 58-63; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Query Match 76.9%; Score 1477.2; DB 2; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAAGAGGTGTTTATCTGTTTATGAGCTGACGAGATTCTGTCGACTGTTCATTA 66
DB 10 AAGAAAGAGGTGTTTATCTGTTTATGAGCTGACGAGATTCTGTCGACTGTTCATTA 69
QY 67 AACATCCCTGGGCTGGATGACGAAGACTTTGATCTGGATTTTAAAGAAATTCAGACA 126
DB 70 AGTAATTCATCTCGAGGTGGTGCAGAGAAATTTGATCTGGATTTTAAAGAAATTCAGACA 129
QY 127 ACAACCGATGTCAGTGTCTTCTCAACAACGACGAACAACAGTGCAGCTGCATTTCTGTGTG 186
DB 130 ACAATCGATGTCAGTGTCTTCTCAACAACGAGGCGACGCTGTGTCTTTCTGTGTG 189
QY 187 GAGTCTGAAATATGTAACCTTTCTTAAAGGATTCGTAAGAAACTTGAACAGTACCGGCA 246
DB 190 GAATCTGAAATATGTAACCTTCCCAAGAGGTTTGCAGAGAAAGACTTGAACAGTACCGGCA 249
QY 247 AATATTAACCTCATATTGTCGAATTCGAATGAGCCCATTTTGAAGAAACAAAGAG 306
|||||

DB 250 AATTAATAACTCCATATTATTCATTCATTCAGTACCAATTTTGAAGAAACAAACAGCAG 309
QY 307 CTAGAGACACTGAGACAAAGATTCTGACATACATCCCTGATTATGATATATTGTTCAG 366
DB 310 CTGAAAAAACAAGGCGCAAGATTCTGACATACATCTGATTAATCTTACATTTGTCAG 369
QY 367 TATGAGGGGGATGTTCAATCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGGAGCCA 426
DB 370 TATGAGGGGGATGTTCAATCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGGAGCCA 429
QY 427 TACTTCCGAAATATCAAAATAGATCCCGCTTTTCACAAAGGCGCATCGAGCTGTGTG 486
DB 430 TATTTGCCATATACGAATATAGTCCCGCTTTTCACAAAGGCGCATCGAGCTGTGTG 489
QY 487 AAAGCTTGGCGCTTGATTCGAAGCAGAACTTAAAGAAATGCAATTAAAGGCGCATCGAG 546
DB 490 AAAGCAGTGGCGCTTGATTCGAAGCAGAACTTAAAGAAATGCAATTAAAGGCGCATCGAG 549
QY 547 GAAATCGCTCAGTACGTAGCAGCAATGACGTCCATTAATTAACGCAAAAGCTGATAT 606
DB 550 CAATTCGACCAATTCGCAATTAAGCAATGATGCTATATATTAACGCAAAAGCTGATAT 609
QY 607 AAGTGATGAATGATGTCAGAGGATTCGAAAGCGATGTGCAAGAGCTAC 666
DB 610 AAGTGATGAATGATGTCAGAGGATTCGAAAGCGATGTGCAAGAGCTAC 669
QY 667 GGTGTTGATGACAAAGGCCGATTTGCGAGTTGCGGATTCGATTCGATTCAGAGAGA 726
DB 670 GGTGTTGATGACAAAGGCCGATTTGCGAGTTGCGGATTCGATTCGATTCAGAGAGA 729
QY 727 AACGACGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 786
DB 730 AATGACGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 789
QY 787 CGAGCAATTAATGCGAATGATTCGAAACGCTGATGATGATGATGATGATGATGATGATG 846
DB 790 CGAGCAATTAATGCGAATGATTCGAAACGCTGATGATGATGATGATGATGATGATGATG 849
QY 847 GGAATATGCGCAAGATTAAGAAATGCGCTCAAGCGAATCTGATTTTCAATTCATC 906
DB 850 GGAATATGCGCAAGATTAAGAAATGCGCTCAAGCGAATCTGATTTTCAATTCATC 909
QY 907 ATGATATGAGTGTGAGGCTTGGAGCTTGGCTTGCATTCGAAACCTTATTCAGCCAA 966
DB 910 ATGATATGAGTGTGAGGCTTGGAGCTTGGCTTGCATTCGAAACCTTATTCAGCCAA 969
QY 967 GCATTCAGTGCAGTGCAGGATTCATTAACATCTCTGAGGAGCGAGCGTGAATGGAGCT 1026
DB 970 GCATTCAGTGCAGTGCAGGATTCATTAACATCTCTGAGGAGCGAGCGTGAATGGAGCT 1029
QY 1027 TACACGACGATTCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
DB 1030 TACACGACGATTCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1089
QY 1087 TTGCGGCTGGGAAATGAAGCGCGAAGCGGATCATCATGTCACCTTGATGAGCTTAA 1146
DB 1090 TTGCGGCTGGGAAATGAAGCGCGAAGCGGATCATCATGTCACCTTGATGAGCTTAA 1149
QY 1147 AACGCAATTAAGTGTGCGCAACGAAACCTGCGCTCAAGCTTGGTTCATTCAGAT 1206
DB 1150 AACGCAATTAAGTGTGCGCAACGAAACCTGCGCTCAAGCTTGGTTCATTCAGAT 1209
QY 1207 AATATTAACAGTGTGACAGTTCCTTCCGCTGCGCGCAAAAGATGAGCGAATCAAG 1266
DB 1210 AATATTAACAGTGTGACAGTTCCTTCCGCTGCGCGCAAAAGATGAGCGAATCAAG 1269
QY 1267 CCGATATGATGAGCGCGCGCAAGTATCAATTTATGAGCAAGATCTTCTTGCACCGCAT 1326
DB 1270 CCGATATGATGAGCGCGCGCAAGTATCAATTTATGAGCAAGATCTTCTTGCACCGCAT 1329
QY 1327 TCTTCTTCTTGGGGAATCATGACGAAATATGCTTATGAGGATGAAAGCTTCATGCA 1386
DB 1330 TCTTCTTCTTGGGGAATCATGACGAAATATGCTTATGAGGATGAAAGCTTCATGCA 1389
|||||

QY 1387 ACACCGATTGTCGGGGAAATGTCACAGCTCCGTGACATTTTGTGAAAAATAGAGA 1446
Db 1390 ACACCGAATGTTGTCGGAAACGTGCACAGCTTCGTGACATTTTGTGAAAAACGAGGC 1449
QY 1447 ATCACTCTTAAGCCTTCCTTATTTGAAGAGCTTTGATTTGAGTGTGATTTGGA 1506
Db 1450 ATCACTCTTAAGCCTTCCTTATTTGAAGAGCTTTGATTTGAGTGTGATTTGGA 1509
QY 1507 TTGGTTATTCGGAACGAAACCAAGAGTGGGCGCAGTGAACCTTGATTAATCGTTGAA 1566
Db 1510 CTGGCTACCGGAAACGTTAACCAAGAGTGGGCGCAGTGAACCTTGATTAATCGTTGAA 1569
QY 1567 GTTGGCTATGGAAGAAATCCAGTCCCTATCAATAGCAAAAGCAATATACCTT 1626
Db 1570 GTTGGCTATGGAAGAAATCCAGTCCCTATCAATAGCAAAAGCAATATACCTT 1629
QY 1627 ACTGCACGCGGCGGCAAGCATTGAAATCTCCCTGTATGTTGCGGATGCCCTGCAAGC 1686
Db 1630 ACTGCTACTGCGCGGCAAGCATTGAAATCTCCCTGTATGTTGCGGATGCCCTGCAAGC 1689
QY 1687 ACTACTGCTTCTGTATACCTGCTGATGATTTGATTTGATTTGATTTGATTTGATTT 1746
Db 1690 ACACTGCTTCTGATACCTGCTGATGATTTGATTTGATTTGATTTGATTTGATTT 1749
QY 1747 ACAAGATATGTCGGGAATGACTTCTCAGACCAATTTGACAAATACCTGGAATGCCGCAAT 1806
Db 1750 ACAAGATATGTCGGGAATGACTTCTCAGACCAATTTGACAAATACCTGGAATGCCGCAAT 1809
QY 1807 AACGTAGAAATGATTTTATTTATTTATTTGATTTGATTTGATTTGATTTGATTT 1866
Db 1810 AACGTAGAAATGATTTTATTTATTTATTTGATTTGATTTGATTTGATTTGATTT 1869
QY 1867 GCATTAATGTCGGCGGTGACCAAACTTCTGTTGGCAATTTGAACTAA 1920
Db 1870 GCATTAATGTCGGCGGTGACCAAACTTCTGTTGGCAATTTGAACTAA 1923

RESULT 3

AAK37279

ID AAK37279 standard; DNA; 1923 BP.

XX AAK37279;
AC 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
XX Bacillus alkaline protease encoding DNA.
XX DE Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX KM washing composition; oxidizing agent; ss.
XX OS Bacillus sp.
XX OS MO9918218-A1.
XX FN 15-APR-1999.
XX PD 07-OCT-1998; 98WO-JP004528.
XX PF 07-OCT-1997; 97JP-00274570.
XX PR
XX (KAO) KAO CORP.
XX PA
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX WPI: 1999-287736/27.
XX DR P-PDB: AAY17091.
XX FT Alkali protease from Bacillus used in washing powders.
XX PS Disclosure; Page 63-68; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1923 BP; 578 A; 417 C; 474 G; 454 T; 0 U; 0 Other;
Query Match 76.9%; Score 1477.2; DB 2; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
QY 7 AAGAAAGAGTGTGTTTATCTGTTTATCACTGACAGGATTCGTGACTGTGATTA 66
Db 10 AAGAAAGAGTGTGTTTATCTGTTTATCACTGACAGGATTCGTGACTGTGATTA 69
QY 67 AACATCCCTGCGCTGCTGATGCAAGACTTTGATTCGATTTTAAAGAAATCAACA 126
Db 70 AGTATCATCTGACAGGCTGGCAAGAAATTTGATCTGATTTTCAAGAAATCAACA 129
QY 127 ACAACGATGTCAGTGGTTTCTCAAAACAGCAAAACAGGCTGGCTGATTTCTGGTG 186
Db 130 ACAACGATGTCAGTGGTTTCTCAAAACAGCAAAACAGGCTGGCTGATTTCTGGTG 189
QY 187 GAGTCTGAAATGTAACCTCTTAAGATGCTTAAGAACTTGAACAGTACCGCA 246
Db 190 GAGTCTGAAATGTAACCTCTTAAGATGCTTAAGAACTTGAACAGTACCGCA 249
QY 247 AATATTAATCTCCATTTTGTGATTTCAATGAGCCCATTTTGAAGAAACAAACAGAG 306
Db 250 AATATTAATCTCCATTTTGTGATTTCAATGAGCCCATTTTGAAGAAACAAACAGAG 309
QY 307 CTAGAGCAACTGAGCAAGAAATCTGCACTACATCCCTGATTAATGATATATTTGTCAG 366
Db 310 CTAGAGCAACTGAGCAAGAAATCTGCACTACATCCCTGATTAATGATATATTTGTCAG 369
QY 367 TATGAGGGGATGTTAGTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGAAGCA 426
Db 370 TATGAGGGGATGTTAGTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTGAAGCA 429
QY 427 TACTTGGCGAATACAAATATGATCCCGAGTTTTCACAAAGGCGATCGACCTGGTG 486
Db 430 TACTTGGCGAATACAAATATGATCCCGAGTTTTCACAAAGGCGATCGACCTGGTG 489
QY 487 AAAGCGTTGGCGCTTGTATGATGAGCAAGCAATTAAGAGTGAATTAAGAGGATCGAG 546
Db 490 AAAGCGTTGGCGCTTGTATGATGAGCAAGCAATTAAGAGTGAATTAAGAGGATCGAG 549
QY 547 GAAATGCTCAGTACGTAAGCAAGCAATGACGTCATTAATTAACGCAAAAGCTGAAATAT 606
Db 550 GAAATGCTCAGTACGTAAGCAAGCAATGACGTCATTAATTAACGCAAAAGCTGAAATAT 609
QY 607 AAGGTGATGAATGATGTCGCAAGAGTATTTGCAAAAGCGATTTGCGACAGAGAGCTTAC 666
Db 610 AAGGTGATGAATGATGTCGCAAGAGTATTTGCAAAAGCGATTTGCGACAGAGAGCTTAC 669
QY 667 GGTGTTATGGAACAAGGCGAATGTCGAGTTCGATTTGATTTGATTTGATTTGATTTGATTT 726
Db 670 GGTGTTATGGAACAAGGCGAATGTCGAGTTCGATTTGATTTGATTTGATTTGATTTGATTT 729
QY 727 AACGACGTTGATGATGATGAAGCTTCGCGGTAAATTAACAGCACTATATGCACTGGGT 786
Db 730 AATGACGTTGATGATGATGAAGCTTCGCGGTAAATTAACAGCACTATATGCACTGGGT 789

```

QY 787 CGGAGGATATATGCGAATGATACGAGCGTACATGTTACCATGTCGAGGTTCCGATTTA 846
DB 790 CGGAGGATATATGCGAATGATACGAGCGTACATGTTACCATGTCGAGGTTCCGATTTA 849
QY 847 GGAATATGCGGCAACGATATGAGGATGAGCTCAAGCGAATCTGGTTTTCATATCATC 906
DB 850 GGAATATGCGGCAACGATATGAGGATGAGCTCAAGCGAATCTGGTTTTCATATCATC 909
QY 907 ATGATATGAGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 966
DB 910 ATGATATGAGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 969
QY 967 GCAATTCAGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1026
DB 970 GCAATTCAGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1029
QY 1027 TACAGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1086
DB 1030 TACAGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1089
QY 1087 TTGCGGCTGCGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1146
DB 1090 TTGCGGCTGCGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1149
QY 1147 AACGCGATTAAGTGGGCGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1206
DB 1150 AACGCGATTAAGTGGGCGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1209
QY 1207 AATATTAACGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1266
DB 1210 AATATTAACGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1269
QY 1267 CCGATGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1326
DB 1270 CCGATGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1329
QY 1327 TCCCTCTTGGGCGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1386
DB 1330 TCCCTCTTGGGCGGATGAGGCTTGGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1389
QY 1387 AACGCGATTAAGTGGGCGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1446
DB 1390 AACGCGATTAAGTGGGCGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1449
QY 1447 ATCACTCTTAAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1506
DB 1450 ATCACTCTTAAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1509
QY 1507 TTGGGTTATCCGAGGAGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1566
DB 1510 TTGGGTTATCCGAGGAGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1569
QY 1567 GTTGCCTTAAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1626
DB 1570 GTTGCCTTAAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1629
QY 1627 ACTGCAAGGCGGAGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1686
DB 1630 ACTGCAAGGCGGAGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1689
QY 1687 ACTGCAAGGCGGAGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1746
DB 1690 ACTGCAAGGCGGAGGCAACGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1749
QY 1747 ACAAGATATGAGGATGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1806
DB 1750 ACAAGATATGAGGATGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1809
QY 1807 AACGTAAGGATGAGGATGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1866
DB 1810 AACGTAAGGATGAGGATGAGGCTTGCCTTCATATGAGGATCTTTATTCAGCGAA 1869

```

```

QY 1867 GCAATTAATGTCGCGGTTGGACCAACAACTTCTGTTGGCAATTTGAACTTA 1920
DB 1870 GCTTATTAACGTAACCGGTTGGACCAACAACTTCTGTTGGCAATTTGAACTTA 1923

RESULT 4
ADM40772
ID ADM40772 standard; DNA; 1923 BP.
XX
AC ADM40772;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding alkaline protease from Bacillus sp. KSM-KP43.
KW alkaline protease; laundry detergent; bleaching agent; detergent;
KW denature-cleaning agent; ds; gene.
OS Bacillus sp.; KSM-KP43.
FH
FT Key Location/Qualifiers
FT CDS 1..1923
FT /tag= b
FT /product= "Bacillus sp. KSM-KP43 alkaline protease"
FT sig_peptide 1..618
FT /tag= a
FT /product= "Alkaline protease prepro sequence"
FT mat_peptide 619..1923
FT /tag= c
FT /product= "Mature alkaline protease"
PN US2004072321-A1.
XX
PD 15-APR-2004.
XX
PF 09-JUN-2003; 2003US-00456479.
XX
PR 26-JUN-2002; 2002JP-00186387.
PR 18-OCT-2002; 2002JP-00304232.
XX
PA (KAOS ) KAO CORP.
PI Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;
DR WPI; 2004-328572/30.
DR P-PSDB; ADM40773, ADM40770, ADM40771.
XX
PT New alkaline protease having a mutant prepro sequence where amino acid
PT residues at positions 52, 75 and 142 are substituted with another amino
PT acid residue, useful as enzyme component of laundry detergents, or
PT bleaching agents.
XX
PS Example 1; SEQ ID NO 3; 29pp; English.
XX
CC The invention relates to an alkaline protease having a prepro sequence.
CC The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
CC acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
CC acid residues at: (a) position 52 is substituted by aspartic acid or
CC arginine; (b) position 75 is substituted by alanine or arginine; and (c)
CC position 142 is substituted by lysine. The alkaline protease is useful as
CC an enzyme which can be incorporated into laundry detergents, bleaching
CC agents, detergents for cleaning hard surfaces or drainpipes, denature-
CC cleaning agents, and detergents for sterilising medical apparatus. The
CC present sequence represents DNA encoding alkaline protease from Bacillus
CC sp. KSM-KP43.
XX
SQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Query Match 76.9%; Score 1477.2; DB 12; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAGAGAGGTCGTTTATCTGTTTATGACGTGACGAGGATTCGTGCACTGTGCACTTA 66

```

Db	10	AAAAAAAAAGGTGTTTATCTGTTTATATAGCTGACGCCATTTTGTGACTGTGGTTA	69
Qy	67	AAACAATCCCTCGGCTGTGTATSCAAGACCTTTTGTATCTGGAATTTTAAAGAAATTCAAACA	126
Db	70	AGTAATTCATCTGACAGGTGTGCAAGGAATTTTGTATCTGTGAATTTCAAGAAATTCAGACA	129
Qy	127	ACAACCGATGTCAGTGGTTTCTTCCAACACGCGACAAACAGTGGCGGCTGCAATTTCTGGTG	186
Db	130	ACAACTGATGTCTAAAGGTGTTCTTCACAGCGGGGACAGACTGGTGGCTGCTCTTTCTGGTG	189
Qy	187	GAGTCTGAAATGTGAAACTTCTTTAAAGATTTGCTTAAAGAACTTGAAACAGTACCGGCA	246
Db	190	GAATCTGAAATGTGAAACTCCCAAAAGGTTTTCAGAAAGACCTTGAAACAGTCCCGGCA	249
Qy	247	AATAATPAACTTCATATTTGTCCAATTCATATGAGCCCACTTTTGAAGAAACAAACAGAG	306
Db	250	AATAATPAACTTCATATTTATCCAAATTCATATGACCAAACTTTTGAAGAAACAAACAGAG	309
Qy	307	CTAGAGACACTGAGAGCAAAAGATTTCCGACTATACCTCGATTAATGACTATATTTGTCAG	366
Db	310	CTGGAAAAAACAAGGGGCAAAAGATTTCCGACTATACCTGATTAATGCTTACATTTGTCAG	369
Qy	367	TATGAGGGGGAGTTCAGTCAAAAGTCCGCTCATTTGAACAAGTGAATACGTGAGACA	426
Db	370	TATGAGGGGGAGTTAAGTCAAGCAACAGCACATTTGACAGCTGAATCCGTGAGACT	429
Qy	427	TACTTGCCGAATATCAAAATAGATCCCGAGCTTTTCACAAAAGCGCACTCGACGTGTG	486
Db	430	TATTTGGCGATATACAAATATGATCCCGAGCTTTTCACAAAAGGGGCACTCAAGCTTGTA	489
Qy	487	AAAGCGTTGGCGCTTGATATCGAAGCAACAAATPAAAGATGCAATTAAGAGCAATCGAG	546
Db	490	AAAGCAATGGCGCTTGATATCAAAAGCAAAAAATTAAGAGGTGCAATTAAGAGGCATCGAA	549
Qy	547	GAATATCGCTCAGTATGACAAAGCAATGACGTCATTAATTAACGGCAAAAGCTGAATAT	606
Db	550	CAAAATCGCACTATTCGCAATPAAAGCAATGATGTGCTATATTAACGGCAAAAGCTGAATAT	609
Qy	607	AAAGTATGAATGATGTGGCCAGAGGTATTTGCAAAAGCGSAGTGTGGCACAGACAGCACTAC	666
Db	610	AAAGTATGAATGATGTGGCCAGAGGTATTTGCAAAAGCGSAGTGTGGCTCAGACAGCACTAC	669
Qy	667	GGTTTGTATGACACAGGCCAGATTGTTCGAGTGGCCGATCTGCAATTTGATATCAGAAAGA	726
Db	670	GGGTGTATGACACAGAGCAGATCTGAGGGTGTGGCCGATACAGGCGCTTGATCAAGTGGC	729
Qy	727	AACGACAGTTGSAATGCAAGAGCCCTTCGCGGCTAAATPAAACGCACTATATGCACTGGGT	786
Db	730	AATGCAAGTTGCATCATGAAAGCCTTCGCGGGGAAATTAATCTGCACTATATATGCACTGGGA	789
Qy	787	CGAGACGAATTAATGCGAATGATPACGAACGGTCAATGTGTACCAATGTGGACAGTTCGGTATTA	846
Db	790	CGAGACGAATTAATGCGAATGATPACGAATGTGTATGCAATGTGGCTGGCTCGGTATTA	849
Qy	847	GAAATATGCGCAACGAATPAAAGAAATGSCACTCAAGCGAATCTGTGTTTCAATCATC	906
Db	850	GAAATATGCGCTCACTPAAATPAAAGAAATGCGCCTCAGGCGAATCTGATCTTCCAATCTATC	909
Qy	907	ATGATATGCAATGTGTGGGCTTGTGAGGCTTGCTCTTCATATTTGCAAACTTATATCAGCCAA	966
Db	910	ATGATATGCAATGTGTGGGGACTTGTGAGAGCACTCACTTGAAATCTTGCAAACTTATATCAGCAA	969
Qy	967	GCATTTACATGACAGTGCACGAATTCATPACAACTCTCGGGGGGCGAGCGGTGAATGGGGCC	1022
Db	970	GCATPACAGTCTGTGGTGCAGAAATTCATPACAACTCTCTGGGAGACAGCACTGAATGGGGCT	1022
Qy	1027	TACACGACAGATTCAGAAATGTGTGATGACTATGTAAAGAAATATGATATGACGATTTCTT	1086
Db	1030	TACACAAACAGATTCAGAAATGTGTGATGACTATGTGCCCAAAATATGATATGACGATTTCTT	1089
Qy	1087	TTTCGGGCTGGGAATGAAAGGCCCAAGCGCGGTACATTCAGTGCACCTGTGTACGGCTTAA	1148

Db	1090	ITCGTGCCGGGAATGAAGGACCGAAGCGGCGAAACATCACTGACACCAAGGACAGCTTAA	1143
Qy	1147	AAAGCCATTAACAGTGGCGCAACCGAAAACTGGCTCCAGCTTCGTTCATGACAT	1206
Db	1150	AATGCAATTAACAGTGGAGCTAACGAAAACTCGGCCAAGCTTTGGGTCTTAAGCGGAC	1209
Qy	1207	AATATTAACCACTTGGCACTTTCTTCCGTGGCCGACAAAAGATGGCGAATCAAG	1266
Db	1210	AATATCAACCACTGGGCACTAGTTCTCTTCAAGTGAACGAAAGATGACGATCAAA	1269
Qy	1267	CCTGATGTCATGGCGCCAGGAGCATATCAATTTATCAGCAAGAATCTTCTCTTGACCCGAT	1326
Db	1270	CCGATGTCATGGCACCGGGAACCTTACTATCAAGCAAGATCTTCTCTTGACCCGAT	1329
Qy	1327	TCCTCTCTTGGGCGAATCATGACAGCAAAATATCTATATGAGTGAACCTCCATGCA	1386
Db	1330	TCCTCTCTTGGGCGAACCATGACAGTAATATGATCATATGAGTGAACCTCCATGCT	1389
Qy	1387	ACACCGAATGTTGGCGGGAATGTTGCACAGCTCCGTGACATTTTGTGAAAAATAGAGA	1446
Db	1390	ACACCGAATGTTGGCTGGAAGAGTGGCACAGCTTGTGTGCAATTTTGTGAAAAACAAGGC	1449
Qy	1447	ATCACTCTTAAGCCCTCCCTATTTGAAGACGCTTGAATGACAGGTGCGCTGATGTTGA	1506
Db	1450	ATCAACCAAAAGCTTCTCTATTAAGACGCACTGATGATGGGTGACGCTGACATGGC	1509
Qy	1507	TTGGGTTATCCGAACCGAAGACCAAGAGTGGGCGGAGTGAACCTCGAATTAATCGTTGAC	1566
Db	1510	CTTGGCTATCCGGAACCGTAACCAAGAGTGGGAGAGATGACATTTGATTAATCCCTGAC	1569
Qy	1567	GTTGGCTATGTTGAACGAATCCAGTGGCCCTATCAACTGACCAAAAGGACATATACCTTT	1626
Db	1570	GTTGGCTATGTTGAACGAATCCAGTGGCTCTATCAACCAAGCAAAAGGACAGTACTCGTTT	1629
Qy	1627	ACTGCAACGGGCGGCAAGCAATGAAAAATCCTCCGTGATGAGTGGAGTGGCCCTGCAAGC	1686
Db	1630	ACTGCTACTGCGGCAAGCTTTTAAAAATCTCCTGGATAGTGTGATGATGCCCTGCAAGC	1689
Qy	1687	ACTACTGCTTCTGTAAACCTGGTCAATGATTTGGATTGGTCAATTAAGCAACCAACGGA	1746
Db	1690	ACAATGCTTCCGTAAGCGCTTGTCAATGATCTGGAACCTTGTCAATTAACGCTCCAAATGGC	1749
Qy	1747	ACAAATATGTTGGGAATGACTTCTCAGCACCAATTTGACAAATTAATGGGAATGGCCGAA	1806
Db	1750	ACACAGTATGTAAGAAATGACTTTTACTTCGCTACATACATGATTAATCTGGAGTGGCCGAAT	1809
Qy	1807	AAAGTAAGAAATGATTTTATTTAATTTGGCCCCAAAGTGAACATTAATCATTTAGAGTGCA	1866
Db	1810	AAAGTAAGAAATGATTTTATTTAATTTGACCAACAAAGCGGAGAGTATACATTTAGAGTACG	1869
Qy	1867	GCATATATATGTCGGTGGACCAAAACCTTCTGTTGGCAATTTGAACTAA 1920	
Db	1870	GCTTATTAACGTACCGGTTGGACCAACACCTTCTGTTGGCAATTTGAACTAA 1923	
RESULT 5			
ADSI4428	ADSI4428 standard; DNA; 1923 BP.		
XX	ADSI4428;		
AC	30-DEC-2004 (first entry)		
DT	Bacillus alkaline protease KP43 gene SEQ ID NO:2.		
DE	protease, enzyme; alkaline protease, laundry detergent, KP43; ds; gene.		
KM	Bacillus sp.; KSM-KP43.		
XX	Key		
OS	location/Qualifiers		
FH	1..1923		
FT	/*C= b		
CDS	/product= "alkaline protease"		
FT			

Db	1510	CTTGCGCTACCCGAAACGGTAAACCAAGATGGGGACAGTGAACATTGGATTAATTCCTGAAC	15059
Qy	1567	GTTGCGCTATGTGAACGAATCCAGTGGCCCTTATCACTAGCCAAAGGACATATACCTTT	16266
Db	1570	GTTGCGCTATGTGAACGAAGTCCAGTTCTGTATCAGCAGCAAAAAGCAGTACGTTT	16299
Qy	1627	ACTGCACGGCGGGCAAGCATTGAAAATCTTCCTGGTATGTCGGAATGCCCTGCAAGC	16866
Db	1630	ACTGCTACTCGCGCGCAACCTTTGGAAAATCTCCCTGGTATGGTCGAAGCCCTGCGAGC	16899
Qy	1687	ACTACTGCTTCGTGAACCTGGTCAATATTTGGATTTGGTCATTACAGCAACAAAGGA	17466
Db	1690	ACAACTGCTTCGTGAACCTGGTCATATATCTGACCTTGTCAATACCGCTTCAAAATGGC	17499
Qy	1747	ACAGATATGTGCGGAATGACTCTTCAGACACCATTGACAATTAACCTGGGAATGGCCGCAAT	18066
Db	1750	ACACAGTATGTAGAGAAATGACTTTTACTTCGCCATACATGATATATCTGGGAATGGCCGCAAT	18099
Qy	1807	AACGTAGAAATGTATTTATTTAATTTCGCCCCCAAAGTGAACATATACATTCGATGAGGTGCAA	18666
Db	1810	AACGTAGAAATGTATTTATTTATTCACACCAAAACGGGACGTAATACATTCGATGAGGTGCAAG	18699
Qy	1867	GCATTAATATGTCGGGTGGAGCAACAAAACCTTCGTTGGCAATGTGAACTTAA	1920
Db	1870	GCTTAATTAACGTACCGGTGGAGCAACAGACCTTCTGTTGGCAATGTGTAATTTAA	1923

Db 919 TAAAGACCTCTTCTAATTGGGCGAACAATTGGAATTTGAAGGAGTAGA 978
Qy 125 CAACAACGATGTCAGTGGTTCTCCAAACAGCGA CAACAGTGGCGTGCATTTCTG 184
Db 979 CACTTACCTTAGAAGGCTGCCAAGCAAGAAAAACGGAAAGGATCTTTCTTG 1038
Qy 185 TGGAGTCTGAAAATGTGAACTTCTTAAAGATTGCTAAAGAACTTGAACAGTACCG 244
Db 1039 TAACTCTGAAAATGTGAAAATCCCAAGAGTATTCAAAAGAACTTGAAGTGTTCAG 1098
Qy 245 CAATATTAATATCCATATTGTCAATTCAATGGCCCATTTTGAAGAAACAAACAGA 304
Db 1089 CGGATTAACAGCTATATATCGTTCAATTTGACGACCTATTTTGAAGAAACGCACTTC 1158
Qy 305 AGCTAGAGCAACTGAGCAAGATTTCTGACTACATCCCTGATTTATGCTATATTGTCG 364
Db 1159 AACTAGAGAGACGGGAGCGAAAATTTCTGATTAATACAGATTAAAGCTTATATTGTCG 1218
Qy 365 AGTATGAGGGGATGTTCAATCAAAAGTCCGCTCATTTGAACAGTGGATCAGTGAAC 424
Db 1219 AATATGATGGGAGTGTAAAGGCGTAACTAAAGCAATTGCGCATTTGGAATCGTGAAC 1278
Qy 425 CATACTTCCGAAATATACAAATATAGATCCCGAGCTTTTCAAAAAGCGCATGACGTCG 484
Db 1279 CATATTTACTTTATATATTAATAGACCGCAATTTATTTCCAGAGAGCTTCTGAATTAG 1338
Qy 485 TGAAGCGTTGGCGCTTATACGAAGCAAAACAATTAAGAAAGTGAATTAAGAGGATCG 544
Db 1339 TAGAAAACAGTACCTTAGATTAAGCAAGAAAGTAAAGAGTACGTTTAAAGAGATTGG 1398
Qy 545 AGGAAATGGCTCAGTACGTAGCAAGCAATGACGTCAATTTATTTACGGCAAGCTGAT 604
Db 1399 AACAAATGGCCCAATACCGCAAAATATGATGATTTATACGTAACCCCAAGCTGAT 1458
Qy 605 ATTAAGTATGATATGATGTCGCAAGGATTTGTCAAAAGCGATGTCAGACAGACGT 664
Db 1459 ACGAAGTTTGAATACGTGGCCCGTGGCATTTGTGAAGAGACAGCTCGCAAAATACCT 1518
Qy 665 ACGGTTTATGACAAAGGCCAGATTGTGCGAGTTGCCGATACGTGATTTGATACAGAA 724
Db 1519 TTGGCTTATGACAAAGGACAGATTGTAGCAGTTGTGCTGATCTGGGCTGATACAGAA 1578
Qy 725 GAAACGATGTCATGATGTAAGGCTTCGCGGTTAAATTAACAGCACTTATATGACCTG 784
Db 1579 GAAATGACAGTTCAGTATGAAAGATTCGCGGTAAGATTAACGCACTATATGACCTG 1638
Qy 785 GTCCGACGATATATGCAATGATACGAACGTCATGATACCATGTCGAGTTGCGTAT 844
Db 1639 GCAAAACGATTAACCCCATGATCCAAATGACATGAAACCATGTTGCTGATCTGTGT 1698
Qy 845 TAGAATATGCGCAACGAATTAAGGATGGCACTCAAGCAATCTGTTTTCAATCCA 904
Db 1699 TAGAATAT---GCTACAAATTAAGGAGATGGCACCGCAAGCAATCTAGCTTTCAATCTA 1755
Qy 905 TCATGGAATAGCAGTGGTGGCTTGGAGGCTTGCTTCAATCTGAAAACCTTATTCAGCC 964
Db 1756 TTATGGAATAGTGTGAGAGGCTGGGAGACTACCTGCTATCTTAACAACTTATTCAGTTC 1815
Qy 965 AAGCATTCAGTGCAGGTCGCAAGATTCATACAACTCTGGGGGGCGACGGTGAATGGGG 1024
Db 1816 AAGCATATATGTCGTGAGCGAAGATTCATACGAATTCATGGGGGGCTCAAGTAAACGTTG 1875
Qy 1025 CCTACACGACAGATTCAGAAATGTGATGACTATGTAAGAAAAATGATATGACGATTC 1084
Db 1876 CCTATACGACAGACTCTGAAATGTTGATGATTAATGTAAGAAAAATGATATGACGATTC 1935
Qy 1085 TTTTTCGGGCTGGGAATGAAGCGCGAACGCGGTACATCAGTGCACCTGTACGGCTA 1144
Db 1936 TTTTTCGGGCTGGGAATGAAGCGACGAGTACGATCAATCAGTGCACGAGAAACAGAA 1995
Qy 1145 AAAACGCAATACAGTGGGCGAACCGAAAACCTGCTCAAGCTTCGATTCATGACG 1204

Db 1996 AAAATGCATTAACAGTTGGGCGAACCGAAAACCTACGTCCAAAGCTTCGATCTTAGCGG 2055
Qy 1205 ATTAATATTAACACAGTGGCAAGTCTCTTCCGTGGCCGCAACAAAGATGGCGAATCA 1264
Db 2056 ATTAATATTAACAGTGGCTCAATCTCTTCAACAGGCTCTACAGATGAGATGAGATTA 2115
Qy 1265 AGCGTATGATGCGCGCAAGGACATPACATTTATACAGAAAGATCTTCTGTCAACCG 1324
Db 2116 AGCGGACGTATGGCACAGATAGTATATCTCTGTGATGATCATTAAGCTCAG 2175
Qy 1325 ATTCCTCTTGGGCGAATCATGACAGAAATATGCTATATGAGTGAACGTCATAG 1384
Db 2176 ATTCCTCATTTGGGCAACATGATATATATGCTACATGGTGTACTTCTATAG 2235
Qy 1385 CAACACGATTTGTCGGGAATGTTGACAGCTCCGTGACATTTTGTGAATAAGAG 1444
Db 2236 CTACTCCAAATGTAGCAGGTATGTGCAATTAAGGAGACATTTGTGAATAAGAG 2295
Qy 1445 GAATCATTCTTAACCTTCCCTATTTGAAGAGCTTTGATTGCAAGTCTGTGATG 1504
Db 2296 GGGTAACTCTTAACCTTCCCTTTTAAAGCTGCTTAAATGTCAGGTGTCGGAATGTTG 2355
Qy 1505 GATTGGTTATCCGAACGAAACCAAGAGATGGGCGAGTGAACCTGGATTAATGTTGA 1564
Db 2356 GACTTGGCTTCCAAATGTATACCAAGATGGGAAAGTAACTGATATTAATCCCTAA 2415
Qy 1565 ACGTTGCTATGTGAACGATTCAGTGCCTATCACTAGCCAAAAGCGACATATACCT 1624
Db 2416 ATGTGCAATTTGTATGAAGAAACGAGCTTATCAACAGTCAAAAAGCAATATTCGT 2475
Qy 1625 TTATGCAACGGCGGGGAGCATTTGAAAATCTCCGTGATATGTCGATGTCGCTGCA 1684
Db 2476 TTACGGCTCAAGCTGTGAACCTTTAAATATATACCTGTGTGTGATGATCACAGTGA 2535
Qy 1685 GCATTAAGCTTCTGTAAACCTGTCATGATTTGATTTGATTTGATTAACAGCAACAAACG 1744
Db 2536 GCAAGAGGATCACTAATCTTAGTAATGATTTAGCTTATGATCACTGACCAATAG 2595
Qy 1745 GAAACAGATATGTCGGGAATGACTTCTCAGACCAATTTGACATTAACCTGGATGGCGCA 1804
Db 2596 GAACATAATATGTCGGAATGACTTTACAGCACGATATGAATTAATGGGATGGCAGAA 2655
Qy 1805 ATTAAGTAAAGATGATTTATTAATTCGCCCCAAAGGGAACATTAACATTAAGGTC 1864
Db 2656 ACAACGTGAAGAAATGTTATATCAATGCTCTCAAGGGAACATTAACGTCGAATGTC 2715
Qy 1865 AAGCATATATGTCGCGTGGACCAACAAACCTCTGTTGGCAATTTGTGAACCTAA 1920
Db 2716 AGGCTTAATATGTAACGATTAAGTCCGAAACCTTTCTTTAGGATTTGATACCTTA 2771

RESULT 8
ADY33779
ID ADY33779 standard; DNA; 1305 BP.
XX
AC ADY33779;
XX
DT 05-MAY-2005 (first entry)
XX
DE Bacillus species alkaline protease coding sequence.
XX
KM mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
XX bleaching detergent; gene; ds.
OS Bacillus sp.
XX
FH Key
FT CDS 1. 1305
FT location/Qualifiers
FT /*tag= a
FT /product= "alkaline protease"
FT /note= "no start codon given"
FT /partial

PN EP1347044-A2.
 XX
 PD 24-SEP-2003.
 XX
 PF 21-MAR-2003; 2P03BP-00006472.
 XX
 PR 22-MAR-2002; 2002JP-00081428.
 XX
 PR 06-JUN-2002; 2002JP-00165987.
 PR 18-OCT-2002; 2002JP-00304230.
 PR 18-OCT-2002; 2002JP-00304231.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
 PI Kobayashi T, Nomura M;
 DR WPI; 2003-846540/79.
 XX P-PSDB; ADY33778.
 XX
 PT New alkaline protease having specific amino acid residue at a specific
 PT position of its amino acid sequence, useful for producing detergent
 PT compositions, laundry detergent, fiber modifiers, leather-treating agents
 PT or pipe cleaners.
 XX
 PS Disclosure; SEQ ID NO 2; 31pp; English.
 XX
 CC The invention relates to an alkaline protease having a fully defined
 CC sequence of 434 amino acids (I) given in the specification, or an amino
 CC acid sequence at least 80% homology with (I), where an amino acid residue
 CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
 CC selected from 16 amino acid residues. The amino acid residues at the
 CC corresponding positions are selected from: position 65: proline; position
 CC 101: asparagine; position 163: histidine, aspartic acid, phenylalanine,
 CC lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or
 CC valine; position 170: valine or leucine; position 171: alanine, glutamic
 CC acid, glycine or threonine; position 273: isoleucine, glycine or
 CC threonine; position 320: phenylalanine, valine, threonine, leucine,
 CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
 CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
 CC acid, arginine or histidine. The alkaline protease is useful for the
 CC production of a detergent composition, such as laundry detergent, fiber
 CC modifiers, leather-treating agents, cosmetic compositions, bath
 CC additives, food modifiers and pharmaceutical compositions. The alkaline
 CC protease may also be used as bleaching detergent, hard surface cleaning
 CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
 CC cleanser for medical tools. The new alkaline protease has a more potent
 CC proteolytic capacity, exhibiting excellent detergency for the removal of
 CC a complex stain, and has high secretion capacity. This sequence
 CC corresponds to the *Bacillus* sp. alkaline protease coding sequence.
 CC
 XX
 SQ Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;
 Query Match 52.2%; Score 1002.6; DB 11; Length 1305;
 Best Local Similarity 85.5%; Pred. No. 1.4e-305;
 Matches 1116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

DB 241 TCACCTAATAAAGAAATGGCGCTCAGCGCAATCTAGTCTTCAATCTATCATGATAGC 300
 QY 916 AGTGTGGGCTTGAAGGCTTGCCTTCCATCTGCAAACTTATTCAGCAAGCAATTCAGT 975
 DB 301 GGTGGGGGACTTGAAGACTTACCTTGAATCTGCAAACTTATTCAGCAAGCAATTCAGT 360
 QY 976 GCAAGTGCAGAAATTCATCAAACTCTTGGGGGCGAGCGGTGAATGGGGCTTACAGACA 1035
 DB 361 GCTGGTGCAGAAATTCATCAAACTCTTGGGGGCGAGCGGTGAATGGGGCTTACAGACA 420
 QY 1036 GATTCAGAAATGTGGATGACTATGTAAGAAATAATGATATGACATCTCTTTTGGGGCT 1095
 DB 421 GATTCAGAAATGTGGATGACTATGTAAGAAATAATGATATGACATCTCTTTTGGGGCT 480
 QY 1096 GGAATGAAAGGCGCAACGCGGTACCATCACTGCACTTGTAAGGCTTAAACCGCATTA 1155
 DB 481 GGAATGAAAGGCGCAACGCGGTACCATCACTGCACTTGTAAGGCTTAAACCGCATTA 540
 QY 1156 ACAATGCGCGCAACGCAAACTCGGTCCAGCTTGGTCTTATGCAATATATTAAC 1215
 DB 541 ACAATGCGCGCAACGCAAACTCGGTCCAGCTTGGTCTTATGCAATATATTAAC 600
 QY 1216 CAGCTTGCAGATTCCTTCCGTTGGGCGCGCAAAAGATGGCGAATCAAGCTGATGTC 1275
 DB 601 CATGTGGCAAGTTCCTTCACTGTGACCGCAAGAGATGAGACGATCAACCGGATGTC 660
 QY 1276 ATGGCGCCAGGACATATTTATCAGCAAGATCTTCTTGCACCCGATTCCTCTTC 1335
 DB 661 ATGGCACGGGGAACGTTCACTATCAGCAAGATCTTCTTGCACCCGATTCCTCTTC 720
 QY 1336 TGGGCAATCATGACAGCAAAATATGCTATATGAGGTGAAGCTCCATGCAACCGATT 1395
 DB 721 TGGGCAACCAAGACAGTAAATATGATCATGAGGTGAAGCTCCATGCAACCGATT 780
 QY 1396 GTTGGGGGAATGTGCACAGCTCGGTGCAATTTTGAATAATGAGAAATCACTCT 1455
 DB 781 GTTGGGAACAGTGTGCACAGCTCTGTGCAATTTTGAATAATGAGAAATCACTCA 840
 QY 1456 AAGCTTCCCTTATTAAGCAAGCACTTGAATGAGGTGCTGATGTTGATGGGTAT 1515
 DB 841 AAGCTTCTCTATTAAGCGGCACTGATGTCGGTGAAGCTGACATCGGCTTGGCTAC 900
 QY 1516 CCGAACGAAACCAAGATGGGCGGAGTACCTTGAATTAATCGTTGAAGCTTGGCTAT 1575
 DB 901 CCGAACGATTAACCAAGATGGGCGGAGTACCTTGAATTAATCGTTGAAGCTTGGCTAT 960
 QY 1576 GTGAACGAATCAGAGCTTATCACTAGCCCAAAAGCAATATACCTTATCTGCAAG 1635
 DB 961 GTGAACGAATCAGAGCTTCTATCAACAGCCCAAAAGCAATATACCTTATCTGCTACT 1020
 QY 1636 GCGGCAAGCACTTGAATAATCTCCCTGATGATGGATGCCCTGCAAGCACTACTGCT 1695
 DB 1021 GCGGCAAGCACTTGAATAATCTCCCTGATGATGGATGCCCTGCAAGCACTACTGCT 1080
 QY 1696 TCTGTAACTTGTCAATGATTTGGATTTGTCATTACAGCAACAAAGCAAGATAT 1755
 DB 1081 TCGTAACTTGTCAATGATTTGACCTTGTCAATTAACGCTCAAAATGCAACAGATAT 1140
 QY 1756 GTGCGGAATGACTTCTCAGCAACCACTTGAATTAATCTGGATGGCGCAATAAGTA 1815
 DB 1141 GTAGGAATGACTTCTCAGCAACCACTTGAATTAATCTGGATGGCGCAATAAGTA 1200
 QY 1816 AATGATTTATTAATTCGCCCAAGTGAACATATTAACATTAAGGTGCAACATATAAT 1875
 DB 1201 AATGATTTATTAATGCAACCAAGCGGAGCTATCAATTAAGGTGCAACATATAAT 1260
 QY 1876 GTGCGGTGGAACCAAACTTCTGTTGGCAATTTGAACATA 1920
 DB 1261 GTACCGGTGGAACCAAGCTTCTGTTGGCAATTTGAACATA 1305

```
ADL25803
ID ADL25803 standard; DNA; 1305 BP.
XX
AC ADL25803;
XX
XX 20-MAY-2004 (first entry)
XX
DE Bacillus alkaline protease coding sequence.
XX
XX Bacillus alkaline protease; washing agent; gene; ds; enzyme.
XX
XX Bacillus sp.
XX
XX Key location/Qualifiers
XX CDS 1..1305
XX /tag=a
XX /partial
XX /product= "Bacillus alkaline protease"
XX /note= "No start codon is given"
XX
XX JP2004008085-A.
XX
XX 15-JAN-2004.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX (KAOS ) KAO CORP.
XX
XX MPI; 2004-094297/10.
XX
XX P-PSDB; ADL25802.
XX
XX Novel mutant alkaline protease produced by substituting the amino acid
XX residues useful as washing agent.
XX
XX Disclosure; SEQ ID NO 2; 21bp; Japanese.
XX
XX The invention comprises a mutant Bacillus alkaline protease which
XX contains substitutions at positions 163, 170 and 434. The mutant alkaline
XX protease is useful as a washing agent. The present DNA sequence encodes a
XX Bacillus alkaline protease of the invention.
XX
XX Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;
XX
XX Query Match 52.2%; Score 1002.6; DB 12; Length 1305;
XX Best Local Similarity 85.5%; Pred. No. 1.4e-305;
XX Matches 116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;
XX
QY 616 AATGATGCGCCAGAGTATGTCTCAACCGATGTGCGACAGACGCTACCGTTTGTAT 675
DB 1 AATATGTGTGCGCGTGAATGTCTCAACCGATGTGCGTCAAGAGCGCTACCGGTTGTAT 60
QY 676 GGAACAAGCCAGATTTGTCGAGTTGCGATACCTGATGTGATACAGAAAGACAGAGT 735
DB 61 GGAACAAGACGATGTGTCGCGTTGCCATACAGGCGTTGATACAGGTGCGCATGACAGT 120
QY 736 TCGATGATGAAGCTTCCGCGGTAAATAACAGCATATATGCACTGGGTGCGAGAT 795
DB 121 TCGATGATGAAGCTTCCGCGGAAATTAATCTGATATATGCACTTGGAGCGAGAT 180
QY 796 AATGCGAATGATGACAGCGGTCTATGTCACCATGTGCGAGGTTCGGTATTAGAAATGCG 855
DB 181 AATGCGCATGATGACAGATGTCTATGTCAGCATGTGCGTCCGTTATGGAACCGC 240
QY 856 GCAAGAAATTAAGAAATGCAAGCTCAAGCGATCTGTTTTCAATCCATCATGATGATGAC 915
DB 241 TCGACTAATTAAGAAATGCGCGCTCAAGCGAATCTAGCTTCCATCTATCATGATGATGAC 300
QY 916 AGTGTGGGCTTGAAGCTTCCATCTCAATCTGCAAACTTATTAGCGCAAGCATTCAGT 975
DB 301 GGTGTGGGCACTTGAAGCACTACCTTCAATCTGCAAACTTATTAGCGCAAGCATTCAGT 360
```

```
QY 976 GCAGGTGCGAAGATTCATACAACTCCCTGGGGGCGAGCGGTGAATGGGGCCCTACAGACA 1035
DB 361 GCTGTGCGAAGATTCATACAACTCTGGGAGCGAGTGAATGGGGCTTACACACA 420
QY 1036 GATTCCAGAAATGTGATGATGATATGTAAGAAAAATGATATGACGATTCCTTTGCGGCT 1095
DB 421 GATTCCAGAAATGTGATGATGATATGTCGCAAAAATGATATGACGATTCCTTTGCGTCC 480
QY 1096 GGGAAATGAAGCCGGAACGGCGGTACCATCTGATGTCACCTGGTACGGCTTAAAAAGCCATA 1155
DB 481 GGGAAATGAAGCCGGAACGGCGGAACCATCTGATGTCACCAAGGACAGCTTAAAAATGCAATTA 540
QY 1156 ACAGTCGGCGCAACCGAAGAACTGTGCTCAAGCTTCGATCTTATGACATATATTAATTAAC 1215
DB 541 ACAGTCGGAGCTTACGAAAGAACTCCGCGCAAGCTTTGGGTCTTATGCGGACATATTAAC 600
QY 1216 CACGTTGCAAGTTCTCTTCCGTCGCGCCGACAAAGATGGGCGAATCAAGCTGATGTC 1275
DB 601 CATGTGGCAGATTCCTTCACTGACGTCGACCGAACCAAGATGAGACGGAATCAACCGGATGTC 660
QY 1276 ATGGCGCGAAGGACATACATTTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCTTC 1335
DB 661 ATGGCACCGGGAAGCTTCACTATCAAGCAAGATCTTCTTTCGACCGGATTCCTCTTC 720
QY 1336 TGGGCGAATCATGACAGAAATATGCTATATGAGGTGAAGCTGCATGCGCAACCGATT 1395
DB 721 TGGGCGAATCATGACAGAAATATGCTATATGAGGTGAAGCTGCATGCGCAACCGATT 780
QY 1396 GTTGCGGGGAATGTTCGACAGCTCCGTGAGCATTTTGTGAAAAATAGAGAAATCACTCT 1455
DB 781 GTTGCTGGAAGCGTGGCAGAGCTTCGTGAGCATTTTGTGAAAAATAGAGGATCAACCA 840
QY 1456 AAGCTTCCCTATTGAAAGCAGCTTTGATTCAGAGTCTGCTGATGTTGATTTGGGTTAT 1515
DB 841 AAGCTTCTCTATTATTAAGCGGCACTGATTCGCGTGAAGCTGATTCGCGCTTGGCTAC 900
QY 1516 CCGAAGCGAAACCAAGATGGGGGCGAGTGAACCTCGATTAATGTTGAAGCTTGCTTAT 1575
DB 901 CCGAAGCGTATCAAGAGATGGGGGCGAGTGAACCTCGATTAATGTTGAAGCTTGCTTAT 960
QY 1576 GTGAACGATTCAGTGCCTTATCACTAGCAGCAAAAGCGACATTAACCTTTATGCAACG 1635
DB 961 GTGAACGATTCAGTTCCTATTCACACAGCAAAAGGAGATGTCGTTTACTGCTACT 1020
QY 1636 GCGGGCAAGCCATTGAAATATCTCCCTGTATGTGCGATGCGCCCTGCAAGCATATGCT 1695
DB 1021 GCGGGCAAGCCATTGAAATATCTCCCTGTATGTGCGATGCGCCCTGCAAGCATATGCT 1080
QY 1696 TCTGTAAACCCGTGCAATGATTTGGATTTGGTCACTTACAGCAAGAAACGAAACAGATAT 1755
DB 1081 TCTGTAAACCCGTGCAATGATTTGGATTTGGTCACTTACAGCAAGAAACGAAACAGATAT 1140
QY 1756 GTCCGGAATGACTTCTGACACATTTGACAAATATCTGGAATGCGCCGCAATATGTAAGA 1815
DB 1141 GTAGGAATGACTTATCTATGCGCATACATATGATTAAGTGGATGCGCCGAATATGTAAGA 1200
QY 1816 AATGATTTATTAATTTGCGCCCAAGTGAACATATATCATTTAGGTGCAAGCATATAT 1875
DB 1201 AATGATTTATTAATTTGCAACCAAGGCGGACGTATCAATTTAGGTGCAAGCTTATATAC 1260
QY 1876 GTGCGGTTGGAACCAAACTTCTGTTGGCAATTTGAACATA 1920
DB 1261 GTACCGGTTGGAACCAAGACTTCTGTTGGCAATTTGAATTA 1305
RESULT 10
AAQ27516
ID AAQ27516 standard; DNA; 1299 BP.
XX
XX AAQ27516;
XX
XX 05-FEB-1993 (first entry)
XX
```

DE Alkali-protease Ya enzyme gene.
XX
KW Alkali resistance; surface active agent resistance; detergency improver;
XX 88.
XX Bacillus sp. Y.
XX
FH Key Location/Qualifiers
FT CDS 1..1299
FT /tag= a
PN JP04197182-A.
XX
XX 16-JUL-1992.
XX PD
XX PF 28-NOV-1990; 90JP-00327110.
XX PR 28-NOV-1990; 90JP-00327110.
XX PA (LIOY) LION CORP.
XX WPI; 1992-288440/35.
XX P-PSDB; AAR26274.
XX
XX DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX resistance and improves detergency.
XX
XX Claim 3; Page 2; 17pp; Japanese.
XX
XX The sequence is that of the alkali-protease Ya enzyme gene which can be
XX used in the recombinant production of Ya enzyme. Ya enzyme is excellent
XX in alkali resistance and surface active agent resistance and improves
XX detergency
XX
XX Sequence 1299 BP; 401 A; 246 C; 297 G; 355 T; 0 U; 0 Other;
SQ
Query Match 38.9%; Score 747.6; DB 2; Length 1299;
Beet Local Similarity 74.0%; Pred. No. 6.5e-225;
Matches 961; Conservative 0; Mismatches 334; Indels 3; Gaps 1;
QY 616 AATGATGTGGCCAGAGTATTGTCAGAGCGAGTGGCAGACAGACGAGTGGTGTAT 675
DB 1 AATGATGTGCAAGAGGAGTATGAAAGCTGATGTCACAAACCAATTACGATTAAT 60
QY 676 GGAAGAAGCCAGATTGTGCAAGTGGCCGATCTGATGATGATGATGATGATGAT 735
DB 61 GGAAGAAGTCACTAGTTCAGTAGCGGACACAGGCTTATGATGATGATGATGAT 120
QY 736 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
DB 121 TCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 796 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
DB 181 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
QY 856 GCAACGAATTAAGAAATGACCTCAAGCGAACTGTGTTTTCATCATCATCATCAT 915
DB 238 GCTTAAATTAAGAAATGACCTCAAGCGAACTGTGTTTTCATCATCATCATCAT 297
QY 916 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
DB 298 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357
QY 976 GCAAGTGCAGAAATTCATCAAACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
DB 358 GCTGAGAGCAAGAAATTCATCAAACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 417
QY 1036 GATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 418 AATCGAGAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 477
QY 1096 GGAATTAAGAAAGCCAGAAAGGCGGTACATCATCATCATCATCATCATCATCAT 1155

DB 478 GGTATTAAGAGTCTTAATTCAGAAACAAATTAATGATGATGATGATGATGATGAT 537
QY 1156 ACAGTCGGGCGCAACCCGAAACCTGCGTCCAGGCTGCTTCTATGATGATGATGAT 1215
DB 538 AGGTGCGGCGCAACCCGAAACCTGCGTCCAGGCTGCTTCTATGATGATGATGATGAT 597
QY 1216 CACGTTGCAAGTTCCTTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1275
DB 598 CATATTCGCAAAATTTTCATGAGAGAGGCTACAGAGGAGTGAAGATTAAGCTGACGTA 657
QY 1276 ATGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335
DB 658 ACAGCTCTGGAACCAATTTATTTATGACAGCTGCTTCTGCTGCTGCTGCTGCTGCT 717
QY 1336 TGGGCGAATCATGACAGCAAAATATGCTTATATGATGATGATGATGATGATGAT 1395
DB 718 TGGGCGAATTTATTAAGTAAATATGCGGTATATGCGGAGTACCTCCATGCGACACTTAT 777
QY 1396 GTTGGCGGGAATGTTGCAAGCTCCGTGACATTTTGTGAAAAATAGAGATCACTCCT 1455
DB 778 GTTGACAGGGAATGTCGCGCAATTAAGTGAAGCAATTTATTAAGATTAATTAATTA 837
QY 1456 AAGCCTTCCCTATTTGAAGCAGCTTGTGATGATGATGATGATGATGATGATGAT 1515
DB 838 AAGCCTTCTTTATTAATTAAGCTGACCTTATGCTGCTGCTGCTGCTGCTGCTGCT 897
QY 1516 CCGAAGCAAAACCAAGATGAGGCGGAGTGAACCTGATTAATGCTTGAAGCTTGTCTAT 1575
DB 898 CTAAGTGTGACCAAGGCTGCGGCGGTGCTTATCTGATTAATGCTTGAATGATGATGAT 957
QY 1576 GTGAACGAATTCAGTGCCTTATCAATGACCAAAAGCAGATATACCTTACTGCAAG 1635
DB 958 GTCAATGAAGCAATGACCTTATGACCAAGCAAAAGCAGATATACCTTACTGCAAG 1017
QY 1636 GCGGCGAAGCCATTTGAAATTCCTCGTATGATGATGATGATGATGATGATGATGAT 1695
DB 1018 GCGGCGAAGCCATTTGAAATTCCTCGTATGATGATGATGATGATGATGATGATGAT 1077
QY 1696 TCTGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755
DB 1078 TCTGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
QY 1756 GTGGGAATGACTTTCAGACACATTTGACAAATACCTGGAATGCGGCAATTAAGTAA 1815
DB 1138 GTAGAAATGATTTATGATTAATCTTATGATTAATCTGGAATGCGGCAATTAAGTAA 1197
QY 1816 AATGATTTATTAATTCGCCCCCAAGTGAACATATACATTAAGTGAAGTGAATTAAT 1875
DB 1198 AATGATTTATTAATTCGCCCCCAAGTGAACATATATTAATTAAGTGAAGTGAATTAAT 1257
QY 1876 GTGCGGTTGACACAAAACTTCTGTTGGCAATGT 1913
DB 1258 GTACCATTCGCGCCACAGCGTTTTCATCACTGATTCGT 1295
RESULT 11
ADM99124
ID ADM99124 standard; DNA; 1332 BP.
XX
XX ADM99124;
XX
XX 15-JUL-2004 (first entry)
XX
XX
XX Bacterial source DNA encoding a protease Segid 7.
XX
XX gene; db; protease; transgenic; detergent; fruit juice; syrup; puree;
KW paper; oral care; contact lens cleaning composition; hairball remedy;
XX antimicrobial.
XX Bacteria.
XX
XX WO2004033668-A2.
XX
XX

Db 709 KBAKSYRBRBMYWYKMGTYRYYRWRSCMTBRAMSRRKRWAGASMSCWYWRGARS 650
QY 814 GGTGATGTAACCATGTGACAGTTGGTATTAAGAAATGGCGCAACGAATTAAGAAATG 873
Db 649 MWYSKSGCAKCCCKTRRYTSSYSTMGYSYSKMSWTSMGMYTCTMTYSMG 590
QY 874 GCACTCAAGCAATCTGTTTTCATCATCATGATGATGAGTGGGCTTGAGGC 933
Db 589 STRBSMGWSGMSRMYRWKMKMRKRYMKYKCTYRMCCTYRMTTTSRSM 530
QY 934 TTGCTTCCAAATCTGCAAACTTATTCAGCCAAAGCATTCAGTGCAGTGCAGAAATTCAT 993
Db 529 YGKRYARTSRKRYRMYKRYKCMYKYGYGMYKSGSYMRVGYCKRCKCYAMCKKAY 470
QY 994 ACAAACTCTGGGGGCGAGCGGTGAATGGGCTTACAG 1032
Db 469 SGMMYMYKYSKMMMSYKYSMMYKYSKMGKAG 431

RESULT 13

ADA71938
ID ADA71938 standard; DNA, 2000 BP.

ADA71938;

DT 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

XX gene; ds.

OS Oryza sativa.

PN WO2003000898-A1.

PD 03-JAN-2003.

PF 22-JUN-2001; 2001WO-1B001105.

PR 22-JUN-2001; 2001WO-1B001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

DR WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

XX Sequence 2000 BP, 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 2.4%; Score 47; DB 8; Length 2000;

Best Local Similarity 11.2%; Pred. No. 0.0033;
Matches 102; Conservative 406; Mismatches 386; Indels 17; Gaps 5;

QY 1009 GCAGCGGTAATGGGGCTTACACAGACAGATTCAGAAATGTGATGATGTAAGAAA 1068
Db 70 SBAERTSGKSSSSGYSKGNKRYKSRKWRGRGRBRBRBRBRBRBRBRBRBRBRBR 129
QY 1069 AATGATATGACATCTTTTGGCGGGTGAATGAAGCCGCAACGGCTTACAT 1128
Db 130 GGGGMMGOKSMYSYMMCYARGCSCKRKSIGSGWGTCTRGARGGSGGSAKYSK 189
QY 1129 GC-ACCTGTGACGCTAAAACGCCATPACAGTGGCGCAACGAAAACCTGCTCAA 1186
Db 190 SMSKMMMSSCGSCGGRSASRYGYSRKYKMYTYSASCMRAYMTYSWACS 249
QY 1187 GCTTGCTTCCATGACATATATTAACAAGTTCAGATGCTTCTCCGCGGCCGA 1246
Db 250 STTWCRSRBRSMWMMKMRKMSRYGYSYKMYMCTAYKYSYSRKCYRGGWRG 309
QY 1247 CAAGAATGAGGCAATCAAGCTGATGATG-----GCGCCAGGACATACATTTTA 1299
Db 310 ATRVWGRGYSMAMMYKMYRGRYKGMKGMWAGRAMMSMCMKSAACYMRBRMRM 369
QY 1300 TCAGCAATCTTCTTCTTGCACCCGATTCCTCTTGGGCGAATCATGACGAATAT 1359
Db 370 TFRRRWAKKSRTSRKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRY 429
QY 1360 GCCTATATGGGTGGAACGTCATGCGAACACCGATTTGGGGGAATGTCACAGCTC 1419
Db 430 GCMTRMSYGMKMYKMSWRRMSKRYKMSRMYR--WRKKCSRTTMMGKTRGMMGTMR 487
QY 1420 CGTGAGCAATTTGTGAAAATAGAGAAATCACTCTTCAAGCTTCTTCAATGAAAGCACT 1479
Db 488 CRYKRSKMKRCRRBR 547
QY 1480 TTGATTCAGGTGCTGCTGATGTTGATGTTGGTTATCCGACGAAACCAAGATGGGC 1539
Db 548 GYVWAGMMKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYKRYK 607
QY 1540 CGAGTGAACCTGGAATAATGTTGAACGTTGCTATGGAACGAATCCAGTCCCATCA 1599
Db 608 SAMSMSRBRKRCRCAISRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBR 667
QY 1600 ACTAGCAAAAAGACATATATCTTCTTCAAGCGGGCGCAACCATTTGAAAATCTCC 1659
Db 668 MYMSKTYAKKGSYWRRYRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBR 727
QY 1660 CTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1719
Db 728 RYMTSMYKCKSKMYRSMYRSMYRSMYRSMYRSMYRSMYRSMYRSMYRSMYRSMYR 787
QY 1720 GATTTGCTATTAACGAC-----CAACGGAACGAATATGTCGGGAATGCTCTGAG 1774
Db 788 YWYMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 847
QY 1775 CACCATTTGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1833
Db 848 ARYMYKATATCATATKATATKATATKATATKATATKATATKATATKATATKATATKAT 907
QY 1834 CCCCAAGTGAACATATATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1893
Db 908 WKARWGSNAVYRMMYKSAKMMYKSGRWGWTYKMYVCTTWKNAACBRATYKMCAGW 967
QY 1894 AACTTCTGCTT 1904
Db 968 AMYSYSWTRY 978

RESULT 14

ACL37108
ID ACL37108 standard; cDNA, 2000 BP.

ACL37108;

DT 02-JUN-2005 (first entry)

DE	Rice stress-regulated promoter SEQ ID NO:15671.
XX	ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW	agriculture.
XX	
OS	Oryza sativa.
XX	
FN	WO2003008540-A2.
PD	
XX	30-JAN-2003.
PF	
XX	21-JUN-2002; 2002WO-US019668.
PR	
XX	22-JUN-2001; 2001US-0300112P.
PR	24-AUG-2001; 2001US-0314662P.
PR	26-SEP-2001; 2001US-0325277P.
PR	21-NOV-2001; 2001US-0332132P.
XX	
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
PI	Moughamer T, Provart N, Rieke D, Zhu T;
XX	
DR	WPI; 2003-248011/24.
XX	
PT	New stress-responsive nucleic acid, useful for altering the
PT	responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT	stress, salt stress or osmotic stress.
XX	
PS	Claim 48; SEQ ID NO 15671; 89pp; English.
XX	
CC	The invention relates to novel abiotic stress responsive polynucleotides
CC	and polypeptides. Also disclosed are vectors, expression cassettes, host
CC	cells, and plants containing such polynucleotides. Also disclosed are
CC	methods for using the polynucleotides and polypeptides to alter the
CC	responsiveness of a plant to abiotic stress. The invention is useful in
CC	agriculture. The nucleic acid is useful for determining whether a test
CC	plant has been exposed to an abiotic stress condition. It is also useful
CC	for selecting an agent that alters abiotic stress regulated
CC	polynucleotide expression in a plant cell, and to identify a homolog or
CC	ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC	molecule and the polypeptide encoded by it are useful in altering the
CC	responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC	stress, osmotic stress or any of their combinations. The present sequence
CC	is used in the exemplification of the invention
SQ	
	Sequence 2000 BP; 510 A; 350 C; 268 G; 518 T; 0 U; 354 Other;
	Query Match 2.1%; Score 40.6; DB 11; Length 2000;
	Best Local Smlarity 11.6%; Pred. No. 0.35;
	Matches 46; Conservative 171; Mismatches 180; Indels 0; Gaps 0
OY	118 ATTCAACAACGAGTGTCTTCCTCACAACGAAACAGTGCGCTGCA 177
DB	44 WYYTCTSKRWMTSMKRYRMYMKWMYKMAAMAYAWCMYAYGYMMKWYSRTWTY 103
OY	178 TTTCGTGTGAGTCTGAAGAATGTGAACCTTTAAGATTGCTAAGAAACTTGAACA 237
DB	104 TRAYKKMKMRBRKMSRBAAAYMYSYMKTSRKMYGYKKRMWRARRYYTWYCMTS 163
OY	238 GTACGGCAATAATTAACCTCATATGTGCCATTCAATGCCCCATTTTAGAAGAAC 297
DB	164 WNRWRYRMYMKRYRAYKMAYTTTTYTTCCKMMWCYGYKMKWAAPASAKRARBARMMNM 223
OY	258 AAACGAGACTGAGACGACCTGAGCAAAAGATTCTGCATACATCCCTGATTATGATAT 357
DB	224 RWKSMTYAATATYSWMSMWATWRRKRWGMWMAVSSTMMRWMMWYCTTRRAMYMKATWT 283
OY	358 ATTTGCGAGTATGAGGGGAGTGTCTAGTCGAAAAAGTCGCTCATGGAACGCGAATCA 417
DB	264 RCTWTMMMAAMTSANRAWCAAGATISATNTMYAKRTTYMKRRRYAYCAMATRTTKWSMN 343
OY	418 GTGAGGCCATCTTCCGCAATAACAAATAGATCCCACCTTTTCACAAAGCGCANTCG 477

Db 344 RITKMMWTMARSGCKKMMMAARNAATTCMAVTTSTYYVYMRRRMRMRMSMAANYSAKYM 403
 Qy 478 ACGCTGATGAAAGCGTTGCGCGCTTGAATACGAAGCA 514
 Db 404 YAKSYYYMYKCAKMKWTTYATYAAAAAKKTRTAAYNGR 440
 RESULT 15
 ID AAS5654 standard; DNA; 522 BP.
 AC AAS5654;
 XX
 DT 19-DEC-2001 (first entry)
 XX
 DE DNA encoding novel subtilisin homologue #56.
 XX
 KM Subtilisin; Savinase; diagnostic; cleaning solution; laundry detergent;
 KM contact lens cleansing solution; dry cleaning; ds.
 XX
 OS Bacillus sp.
 XX
 PN MO2001.75087-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001MO-US010781.
 XX
 PR 03-APR-2000; 2000US-0194143P.
 XX
 PA (MAXY-) MAXYGEN INC.
 XX
 PI Nees JE, Welch M, Giver LD, Cherry JR, Borchert TV, Stemmer WPC;
 PI Marshull J;
 DR WPI, 2001-616689/71.
 XX
 DR P-PSDB; AAU38799.
 XX
 PT Novel subtilisin homologue polypeptides having improved endo-protease
 PT activity relative to mature Savinase subtilisin polypeptide, useful as
 PT component of cleaning solutions e.g. laundry detergents, dry cleaning.
 XX
 PS Claim 15; Page 82; 138pp; English.
 XX
 CC The invention relates to an isolated polypeptide with improved endo-
 CC protease activity relative to subtilisin homologue polypeptide, Savinase
 CC (RPM). The invention also relates to an integrated system comprising a
 CC computer or computer readable medium comprising a database comprising one
 CC or more sequence records. Each record comprises one or more character
 CC string corresponding to a nucleic acid or protein sequence of AAS56599-
 CC AAS56728 or AAU38744-AAU38873 respectively. The integrated system
 CC comprises a user input interface allowing a user to selectively one or
 CC more sequence record. The integrated system is useful for presenting
 CC information pertaining to one of several sequence records stored in a
 CC database. The method involves determining a list of one or more character
 CC strings corresponding to the sequence or its subsequence, determining
 CC which character strings of the list are selected by a user, and
 CC displaying the selected character strings or aligning the selected
 CC character string with an additional character string. The method further
 CC involves displaying an alignment of the selected character string with
 CC the additional character string, and displaying the list. The
 CC polynucleotides are useful as probes, primers, sense and antisense
 CC nucleotides. The subtilisin homologue polypeptides are useful for
 CC producing antipodes which have diagnostic uses related to the activity,
 CC distribution and expression of subtilisin homologues. The subtilisin
 CC homologue polypeptides are used in compositions that serve as cleaning
 CC solutions in a wide variety of applications including laundry detergents,
 CC contact lens cleansing solutions, and dry cleaning. AAS56599-AAS56728
 CC represent the coding sequences of novel subtilisin homologues of the
 CC invention
 XX
 S0 Sequence 522 BP; 143 A; 124 C; 141G; 114 T; 0 U; 0 Other;

```

Query Match      2.1%; Score 40; DB 4; Length 522;
Best Local Similarity 72.2%; Pred. No. 0.25;
Matches 52; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy      1343 ATCATGACAGCAATATGCTTATATGGTGAACGTCCATGGCAACCCGATTGTGCGG 1402
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      447  ATCCGGGCGAGCATATGCTTCAATGAGCGGCGACATCAATGGCAACCCGATTGTGCGG 506
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      1403 GGAATGTTGCAC 1414
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      507  GCGTGCAGCAC 518
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    
```

Search completed: April 8, 2006, 03:01:47
 Job time : 2983.67 secs

Y:

Y:

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 11 8, 2006, 01:05:15 ; Search time 6935.08 Seconds
(without alignments)
12953.156 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atggaagaagaagaagtcgtc.....cgttcgaactgtgaactaa 1920

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

BST:
1: gb_eat1:*
2: gb_eat2:*
3: gb_eat3:*
4: gb_hic:*
5: gb_eat4:*
6: gb_eat5:*
7: gb_eat6:*
8: gb_eat7:*
9: gb_gss81:*
10: gb_gss82:*
11: gb_gss83:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	47.8	2.5	997	10	CNS005TR
2	45	2.3	449	3	AL060767 Drosophila
3	43.6	2.3	633	3	BJ359939 BJ359939
4	43.6	2.3	641	3	BJ369190 BJ369190
5	43.4	2.3	494	3	BJ393925 BJ393925
6	43.4	2.3	601	3	BJ334651 BJ334651
7	43	2.2	725	8	DR166768 RTRP081_1
8	41.6	2.2	634	5	BK253415 BK253415
9	41	2.1	672	8	DN732737 DN732737
10	41	2.1	1265	8	DN679903 DN679903
11	41	2.1	1358	8	DN689766 CGX66-E11
12	40.8	2.1	678	10	CNS005TR
13	40.4	2.1	854	8	DR951879 EST114341
14	39.8	2.1	426	5	AZ145773 SP_0019_B
15	39.8	2.1	872	7	CO116672 GR_B019
16	39.8	2.1	1101	10	CNS0039G
17	39.6	2.1	560	5	BQ753317 BQ753317
18	39.6	2.1	579	3	BM442028 EBAN01_SQ
19	39.6	2.1	696	3	BM442170 EBAN01_SQ
20	39.6	2.1	700	2	BF574449 BF574449
21	39.6	2.1	710	5	BQ764471 EBAN01_SQ
22	39.6	2.1	721	10	CZ280243 cp10d04.r

c 23	39.6	2.1	737	5	BQ764315
c 24	39.6	2.1	929	6	CA973725
c 25	39.6	2.1	1211	4	BC029393
c 26	39.6	2.1	1577	4	BC012590
c 27	39.4	2.1	1030	8	DN703987
c 28	39.4	2.1	1234	8	DN709335
c 29	39.2	2.0	466	3	BM442304
c 30	39.2	2.0	559	10	CM049955
c 31	39.2	2.0	599	10	CM112692
c 32	39.2	2.0	600	9	BZ360416
c 33	39.2	2.0	608	9	BZ364026
c 34	39	2.0	936	10	CL109289
c 35	38.8	2.0	695	1	A1729280
c 36	38.8	2.0	897	10	CL102311
c 37	38.8	2.0	1101	10	CNS0037Q
c 38	38.6	2.0	423	9	AQ437507
c 39	38.6	2.0	451	9	A2745756
c 40	38.6	2.0	655	9	BH082837
c 41	38.6	2.0	717	3	B1335636
c 42	38.4	2.0	629	10	CM118700
c 43	38.4	2.0	634	10	CE723448
c 44	38.4	2.0	690	3	B1698380
c 45	38.4	2.0	732	8	DN201554

ALIGNMENTS

RESULT 1
CNS005TR 997 bp DNA linear GSS 03-JUN-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL060767.1 GI:4943573
VERSION
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mosmoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1. 997
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12K22"
/clone_id="RPCI-98"
/note="end : TET3"

ORIGIN

	Query Match	2.5%	Score 47.8;	DB 10;	Length 997;
	Best Local Similarity	18.5%;	Pred. NO. 0.035;		
	Matches 74;	Conservative 142;	Mismatches 185;	Indels 0;	Gaps 0;
Oy	483	GGTGAAACCGCTGGCGCTTGATGCAGACAGAACATTAAGAAGTGCATTAAAGGCACT	542		
Dd	927	RRARRRRRGRABRAGRGRSAGGGRGGRBAAAABAARDARBARABARBRRAAATA	868		
Oy	543	CGAGGAATCCCTCAGTACGTACGACGAACAATGACTCCATTATATTACGCCAAAGCTCGA	602		
Dd	867	RARGARRRRRRRRRGARRRRARAGAPAGRRGRRRRRRARRRRGAARARRRGRA	808		
Oy	603	ATATTAGCTGATGAAATGATGTGGCAGAGCTATTTCACAGCGGATGTGCCACAGCAG	662		
Dd	807	RAARBARRRRAGRRRRRRARGARRRRRGRRRRGRRRRRRAARARARARARAARA	748		
Oy	663	CTACGGTTTGTATGACAAAGCCAGATTGTGCAAGTCCCATCTGATTCGATTTGGATACAG	722		
Dd	747	AARRRRARAGAAARRRRRRRRRRRRRAGAAARRRRARARAGAAARRRRRRGR	688		
Oy	723	AAGAAGCACGTTGCATGCATGAAGCCTCCGCGGTAAATTAAGACACTATANGCACT	782		
Dd	687	GGAARRRRRGRRRRRGAARRRRRRRRRTTARRRRRRRRKAGAAARRBARGARRRRRGR	628		
Oy	783	GGGTCGACGAATATGCGAATGATACCAACGTCATCGTACCCTATGTGGACGTTCCGCT	842		
Dd	627	RGRRRRGRARGBRRRAGRRRRGRARAGRRRRARAGRRRRRRBARAAGARRGRRRRRA	568		
Oy	843	ATTAGGAATGCCGACGAATTAAGGAATGGCACTTCAG	883		
Dd	567	RRARGAARRRRRRARRRRRRRAARRAGARAGARAGARAG	527		

RESULT 2	
BJ359939	
LOCUS	BJ359939 449 bp mRNA linear EST 07-MAR-2002
DEFINITION	BJ359939 Dictyostelium discoideum cDNA library, CF Dictyostelium
ACCESSION	dictoideum cDNA clone ddc3k09 5', mRNA sequence.
VERSION	BJ359939
KEYWORDS	Dictyostelium
SOURCE	Dictyostelium discoideum
ORGANISM	Dictyostelium discoideum
REFERENCE	Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
AUTHORS	1 (bases 1 to 449)
TITLE	Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T. Full length cDNA of Dictyostelium discoideum at the culmination stage
JOURNAL	Unpublished (2002)
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@gene.nig.ac.jp.
FEATURES	location/genes,nig.ac.jp.
source	1..449 location/Qualifiers
	/organism="Dictyostelium discoideum"
	/mol_type="mRNA"
	/strain="AX4"
	/db_xref="taxon:44689"
	/clone="ddc3k09"
	/sex="mat A"
	/dev_stage="Culmination stage"
	/clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN	

Query Match 2.3%; Score 45; DB 3; Length 449;
 Best Local Similarity 61.5%; Pred. No. 0.18;
 Matches 72; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Db 87 ATACATGAGATATATTTTGTTCATCTTCATCAAAAGTCCACACATGATGTAAGAT 146
 Qy 1363 CAAGCTGATGTCATGGGCCAGGACATACATTTTATCAGCAAGATCTTCTTGGC 1319
 Db 147 GAAACGTATTTAGTGGCCCGTGGATATATTTTTCATCGGCAATCTTAATGATGTC 203

RESULT 3				
BJ369190				
LOCUS				
DEFINITION	BJ369190	633 bp	mRNA	linear EST 08-MAR-200
	BJ369190 Dictyostelium discoideum	discoideum cDNA library, CF		Dictyostelium
	Dictyostelium cDNA clone ddc49116 5'			cDNA sequence.
ACCESSION	BJ369190			
VERSION	BJ369190..1	GI:19278573		
KEYWORDS	EST.			
SOURCE	Dictyostelium discoideum			
ORGANISM	Dictyostelium discoideum			
	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.			
REFERENCE	1 (bases 1 to 633)			
AUTHORS	Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.			
TITLE	Full length cDNA of Dictyostelium discoideum at the culmination stage			
JOURNAL	Unpublished (2002)			
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yatae, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.			

Query Match	2.3%	Score 43.6;	DB 3;	Length 633;
Best Local Similarity	61.4%;	Pred. No. 0.54;		
Matches	70;	Conservative 0;	Mismatches 44;	Indels 0;
			Gaps	0
Oy	1205	ATATATTTAAACGAGTTGCAAGTTCCTTCCGCGGCCGACAAAGATGGCGAATCA	126	
Db	376	ATATGAAATATATATGGATCATCTTCATCAAAAGGTCGACACATATGGTAGATTGA	435	
Oy	1265	AGCCTGATGTCATGGCGCCAGGGAATACATTTTATACGAAAGATCTTCTCTTG	1318	
Db	436	AACCTGATATGTTGCACCTGGTGAATTTATTACATCGCAAGATCAATGGTG	489	

Query Match	2.3%	Score 43.6	DB 3	Length 633
Best Local Similarity	61.4%	Pred. No. 0.54		
Matches	70	Conservative	0	Mismatches 44; Indels 0; Gaps 0
Oy	1205	ATATATATTAACCAAGTTCGACAGTCTCTTCCGTCGCGCCGACCAAAAGATGGCGCATCA	1266	
Db	376	ATATATGAATAATATATATGGATCATTTCTCATCAAAAGGCTCAACATCATATGGTAAATTGA	435	
Oy	1265	AGCCTGATGTCATGGCCGACGGAACATATCATTTTATACGCAAGATCTTCTCTTG	1318	
Db	436	AACCTGATATATGTTGACACTGATGTAATATATTACATCGGCAAGATCAATGTG	489	
RESULT 4				
Bj393925				
LOCUS				
DEFINITION	Bj393925 Dictyostelium discoidium cDNA library, SF Dictyostelium			
ACCESSION	Bj393925			
VERSION	Bj393925.1			
KEYWORDS	GI:19305011			
SOURCE	EST.			
ORGANISM	Dictyostelium discoidium			
REFERENCE	Dictyostelium discoidium			
AUTHORS	Bukaryota; Mycetozoa; Dictyostelida; Dictyostelium.			
TITLE	1 (bases 1 to 641)			
JOURNAL	Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.			
COMMENT	Full length cDNA of Dictyostelium discoidium at the slug stage			
	Unpublished (2002)			
	Contact: Tadashi Shin-i			
	Center For Genetic Resource Information			

National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1. 641
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dd83105"
/sex="mat A"
/dev_stage="Slug stage"
/clone_1ib="Dictyostellium discoideum cDNA library, SF"

ORIGIN

Query Match 2.3%; Score 43.6; DB 3; Length 641;
Best Local Similarity 61.4%; Pred. No. 0.54; Indels 0; Gaps 0;

Matches 70; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1205 ATATATATTAACCAAGTTGACAGTTCTCTCCGTCGCCGACAAAGATGGCGAATCA 1264

DB 169 ATATGAAATATATATGATGATCTTCATCAAGGTCACACATGATGTATGATGA 228

QY 1265 AGCCTGATGTCAGGGCCGACGACATACATTTTATCAGCAAGATCTTCTTGG 1318

DB 229 AACCTGATATGATGACCTGTGGAATATATATCATCGGCAAGATCAATGCTG 282

RESULT 5

LOCUS

BU334651 494 bp mRNA linear EST 05-MAR-2002
Dictyostellium discoideum cDNA library, AF Dictyostellium

discoideum cDNA clone dda47f08 5', mRNA sequence.

ACCESSION BU334651 GI:19164781

VERSION

KEYWORDS Dictyostellium discoideum

SOURCE

Dictyostellium discoideum

ORGANISM

Dictyostellium discoideum

REFERENCE

Bukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

AUTHORS

Unruhshara, H., Tanaka, Y., Kohara, Y. and Shin-1, T.

TITLES

Full length cDNA of Dictyostellium discoideum at the aggregation

JOURNAL

stage

COMMENT

Unpublished (2002)
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1. 494
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda47f08"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_1ib="Dictyostellium discoideum cDNA library, AF"

ORIGIN

Query Match 2.3%; Score 43.4; DB 3; Length 494;
Best Local Similarity 60.7%; Pred. No. 0.57; Indels 0; Gaps 0;

Matches 71; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1203 AGATATATTAACCAAGTTGACAGTTCTCTCCGTCGCCGACAAAGATGGCGAAT 1262

DB 244 ATACATATGAAATATATATTTGTTTCATCTTCATCAAAAGGTCACACATGATGTATGAT 303

QY 1263 CAAGCTGATGTCATGGCCGACGACATACATTTTATCAGCAAGATCTTCTTGC 1319

DB 304 GAAACCTGATTAATGTTGCCCTGTGGAATATATTAATCAATGCGCAAGATCAATGCTGC 360

RESULT 6

LOCUS

BU387574 601 bp mRNA linear EST 08-MAR-2002
Dictyostellium discoideum cDNA library, SF Dictyostellium

discoideum cDNA clone dda3a18 5', mRNA sequence.

ACCESSION

BU387574

VERSION

BU387574.1 GI:19296958

KEYWORDS

Dictyostellium discoideum

SOURCE

Dictyostellium discoideum

ORGANISM

Bukaryota; Mycetozoa; Dictyostellida; Dictyostellium.

REFERENCE

1 (bases 1 to 601)

AUTHORS

Unruhshara, H., Tanaka, Y., Kohara, Y. and Shin-1, T.

TITLES

Full length cDNA of Dictyostellium discoideum at the slug stage

JOURNAL

Unpublished (2002)

COMMENT

Contact: Tadaeu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1. 601
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda3a18"
/sex="mat A"
/dev_stage="Slug stage"
/clone_1ib="Dictyostellium discoideum cDNA library, SF"

ORIGIN

Query Match 2.3%; Score 43.4; DB 3; Length 601;
Best Local Similarity 60.7%; Pred. No. 0.61; Indels 0; Gaps 0;

Matches 71; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1203 AGATATATTAACCAAGTTGACAGTTCTCTCCGTCGCCGACAAAGATGGCGAAT 1262

DB 47 ATACATATGAAATATATATTTGTTTCATCTTCATCAAAAGGTCACACATGATGTATGAT 106

QY 1263 CAAGCTGATGTCAGGGCCGACGACATACATTTTATCAGCAAGATCTTCTTGC 1319

DB 107 GAAACCTGATTAATGTTGCCCTGTGGAATATATTAATCAATGCGCAAGATCAATGCTGC 163

RESULT 7

LOCUS

DR166768/c 725 bp mRNA linear EST 20-JUN-2005
RTPHOS1_14_G08_b1_A029 roots minus phosphorous Pinus taeda cDNA

clone RTPHOS1_14_G08_A029 3', mRNA sequence.

ACCESSION DR166768

VERSION DR166768.1 GI:68026324

KEYWORDS EST.

Pinus taeda (loblolly pine)

SOURCE Pinus taeda

ORGANISM Pinus taeda

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 725)

Pratt, L., Cordomier-Pratt, M.M., Lorenz, W.W., Zimmermann, C. and

Dean, J.F.D.

An EST database from phosphorous-deficient loblolly pine (Pinus

taeda) roots

Unpublished (2005)

Other ESTs: RPHOS1_14_G08_GL_A029

Contact: Cordomier-Pratt MM

Laboratory for Genomics and Bioinformatics

ORIGIN

Query Match 2.3%; Score 43.4; DB 3; Length 601;
Best Local Similarity 60.7%; Pred. No. 0.61; Indels 0; Gaps 0;

Matches 71; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1203 AGATATATTAACCAAGTTGACAGTTCTCTCCGTCGCCGACAAAGATGGCGAAT 1262

DB 47 ATACATATGAAATATATATTTGTTTCATCTTCATCAAAAGGTCACACATGATGTATGAT 106

QY 1263 CAAGCTGATGTCAGGGCCGACGACATACATTTTATCAGCAAGATCTTCTTGC 1319

DB 107 GAAACCTGATTAATGTTGCCCTGTGGAATATATTAATCAATGCGCAAGATCAATGCTGC 163

QY 1203 AGATATATTAACCAAGTTGACAGTTCTCTCCGTCGCCGACAAAGATGGCGAAT 1262

DB 244 ATACATATGAAATATATATTTGTTTCATCTTCATCAAAAGGTCACACATGATGTATGAT 303

QY 1203 AGATATATTAACCAAGTTGACAGTTCTCTCCGTCGCCGACAAAGATGGCGAAT 1262

DB 244 ATACATATGAAATATATATTTGTTTCATCTTCATCAAAAGGTCACACATGATGTATGAT 303

The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

Email: mmprat@uga.edu
RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia), plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FIRC) and the CCLONES project a the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: M13-21 (GTAAACGACGCGCAG)
POLYA=No.

FEATURES
source

Location/Qualifiers
1..725
/organism="Pinus taeda"
/mol_type="rRNA"
/strain="3 CCLONES"
/db_xref="taxon:3352"
/clone="RTPHOS1.14.G08.A029"
/lab_host="PH10B-T1 phage-resilient E. coli"
/clone_lib="Roots minus phosphorous"
/note="Organ: Root; Vector: pSL1180; Site:1: EcORI; Site:2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 139 days (July 28 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Starting five weeks prior to harvesting roots for mRNA preparation, the weekly fertilizer regimen for the potted trees was modified to omit phosphorus from the 0.5x Hoagland's solution (5 applications). For two days prior to starting the modified fertilizer regimen, pots were flushed extensively with water to reduce residual levels of phosphorus. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 2.2%; Score 43; DB 8; Length 725;
Best Local Similarity 49.8%; Pred. No. 0.85;
Matches 109; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Gy 77 CGGCTGGTGAAGCAAGCACTTTGATCTGGATTTTAAAGAAATTCAACACACACCGATG 136
Db 510 CGGCTGATCTGGTGAAGCTGCTGATCGCAACGTCATGCAAAAAGACCAACCGTGG 451
Gy 137 TCAGTGGTTTCTCCAAACAGCAACAGTGGCGCTGATTTCTGGTGAAGTCTGAAA 196
Db 450 AAACAATCAAGACACACACTTAAGCATCTCGAGCTTGCATGGAATGGCTGCAA 391
Gy 197 ATGTGAAACTTCTTAAAGATTGCTAAAGAACTTGAACAGTACCGCAATTAATAAC 256
Db 390 AAGTTAAGCTTCTTACGTGAGTGAAGACATATTAAGCAGACCTTGATGGCCAAAG 331
Gy 257 TCCATATTGTCATTAATGAGCCCATTTTGAAGAAA 295
Db 330 AGCGTTGACACATTAAGAGGAAATTAAGTACTAA 292

RESULT 8
BX253415 634 bp mRNA linear EST 24-FEB-2003
LOCUS BX253415 Pinus pinaster differentiating xylem adult Pinus pinaster
DEFINITION cDNA clone PB083C05, mRNA sequence.
ACCESSION BX253415

VERSION BX253415.1 GI:28520153
KEYWORDS EST.
SOURCE Pinus pinaster
ORGANISM Pinus pinaster

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Caton, F.R., Le Provost, G., Garcia, V., Barre, A., Frigerio, J.-M.,
Palva, J., Fereireiro, P., Avila, C., Kourte, J.-F., Brach, J., de
Dauvergne, A., Canovas, F.M. and Plomion, C.
Transcriptome analysis of wood formation in maritime pine
unpublished (2003)
Contact: Frigerio JM
Genetique et Amelioration 69

FEATURES
source

INRA
route d'Arcachon 33612 Cestas CEDEX France
Email: Frigerio@piereton.inra.fr
Email: Frigerio@piereton.inra.fr.

Location/Qualifiers
1..634
/organism="Pinus pinaster"
/mol_type="rRNA"
/strain="ecotype: Corsican"
/db_xref="taxon:71647"
/clone="PB083C05"
/issue_type="differentiating xylem"
/dev_stage="adult"
/clone_lib="Pinus pinaster differentiating xylem adult"
/note="Vector: Uni-Zap XR lambda (Stratagene); Site:1: Eco RI; Site:2: Xho I; A composite cDNA library was made with mRNA isolated from normal, compression, opposite, early and late wood of Maritime pine uni-directionally cloned into Uni-ZAP XR using the Zap-CDNA Synthesis kit (Stratagene). pBlueScript SK(-) plasmids were obtained by in vivo mass excision. The nucleotide sequence of the 5' end was obtained by automated sequencing with the T3 primer by GENOME EXPRESS, Meylan, France"

ORIGIN

Query Match 2.2%; Score 41.6; DB 5; Length 634;
Best Local Similarity 49.1%; Pred. No. 2.1;
Matches 110; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Gy 77 CGGCTGGTGAAGCAAGCACTTTGATCTGGATTTTAAAGAAATTCAACACACCGATG 136
Db 81 CAGCTGATCTGGTGAAGCTGCTGATCGCAACGTCATGCAAAAAGACCAACCTGG 140
Gy 137 TCAGTGGTTTCTCCAAACAGCAACAGTGGCGCTGATTTCTGGTGAAGTCTGAAA 196
Db 141 AAACAATCAAGACACACACTTAAGCATCTCGAGCTTGCATGGAATGGCTGCAA 200
Gy 197 ATGTGAAACTTCTTAAAGATTGCTAAAGAACTTGAACAGTACCGCAATTAATAAC 256
Db 201 AAGTTAAGCTTCTTACGTGAGTGAAGACATTAAGCAGACCTTGATGGCCAAAG 260
Gy 257 TCCATATTGTCATTAATGAGCCCATTTTGAAGAAA 300
Db 261 AGCGTTGACACATTAAGAGGAAATTAAGTACTTAAGGAAA 304

RESULT 9
DN732737 672 bp mRNA linear EST 31-MAR-2005
LOCUS DN732737
DEFINITION CNB66-A05.x1d-t SHG-CNB Gasterosteus aculeatus cDNA clone
ACCESSION DN732737
VERSION DN732737.1 GI:62108973
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;

REFERENCE
1 (bases 1 to 672)
Kingsley,D.M., Petchel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schmutz,J. and Myers,R.M.
Expressed sequence tags from *Gasterosteus aculeatus*
Unpublished (2003)
JOURNAL
COMMENT
Stanford Human Genome Center
Contact: Grimwood, Jane
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgc.stanford.edu
Plate: 66
High quality sequence start: 23
High quality sequence stop: 672.
Location/Qualifiers

FEATURES
SOURCE

1..672
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CNC66-A05"
/sex="mixed male and female"
/tissue_type="brain"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="SHGC-CNB"
/note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligot sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTCTAGATCGCGAGCGCGCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_construction_fa.q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems:
<http://www.openbiosystems.com/stickleback>"

ORIGIN

Query Match 2.1%; Score 41; DB 8; Length 672;
Best Local Similarity 49.8%; Pred. No. 3.3;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
85 GATGCAAGACTTTGATCTGGATTTAAGGAATTCAACAAACCGATGTCAGTGT 144
Db GAGACACGAGTCTCATTTTCATCATCAGTGTCTTTCATTCAAAAAGAGGCCCTCT 317
145 TTCTCAACAGCAGCAACAGTGCCTGTCATTTCTGTGAGTCTGAAATGTGAAA 204
Db GCCCCCCCCCCCAAGCAAGCAACTATTAAGTCAACAAGTTCACACAAAGTATTA 377
318 GCCCCCCCCCCCAAGCAAGCAACTATTAAGTCAACAAGTTCACACAAAGTATTA 377
205 TTCTTAAGAGTGTCTAAAGAACTTGAACAGTACCGCAATATTAATCTCATATT 264
Db TTACATCAACAACATCCCAAGTATATAGAAATGATACAAAGCTTTAAGCTTACATTG 437
378 TTACATCAACAACATCCCAAGTATATAGAAATGATACAAAGCTTTAAGCTTACATTG 437
265 GTCCCAATCAATGAGCCCATTTTGAAGA 293
Db GTGCATATGGTGCCTCTTTAAGAGA 466

RESULT 10
DN679903

LOCUS
DN679903 1265 bp mRNA linear EST 30-MAR-2005
DEFINITION
CGX08-G08.x1d-t-SHGC-CGX *Gasterosteus aculeatus* cDNA clone
CGX08-G08 3', mRNA sequence.
ACCESSION
DN679903
VERSION
DN679903.1 GI:62022153
KEYWORDS
EST.
SOURCE
ORGANISM
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; *Gasterosteiformes*;
Gasterosteidae; *Gasterosteus*.
1 (bases 1 to 1265)
REFERENCE
Kingsley,D.M., Petchel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schmutz,J. and Myers,R.M.
Expressed sequence tags from *Gasterosteus aculeatus*
Unpublished (2003)
JOURNAL
COMMENT
Stanford Human Genome Center
Contact: Grimwood, Jane
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@hgc.stanford.edu
Plate: 08
High quality sequence start: 15
High quality sequence stop: 593.
Location/Qualifiers

FEATURES
SOURCE

1..1265
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CGX08-G08"
/sex="mixed male and female"
/tissue_type="eyes"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_11b="SHGC-CGX"
/note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligot sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTCTAGATCGCGAGCGCGCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna_library_construction_fa.q.php#8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems:
<http://www.openbiosystems.com/stickleback>"

ORIGIN

Query Match 2.1%; Score 41; DB 8; Length 1265;
Best Local Similarity 49.8%; Pred. No. 4;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
85 GATGCAAGACTTTGATCTGGATTTAAGGAATTCAACAAACCGATGTCAGTGT 144
Db GAGACACGAGTCTCATTTTCATCATCAGTGTCTTTCATTCAAAAAGAGGCCCTCT 263
204 GAGACACGAGTCTCATTTTCATCATCAGTGTCTTTCATTCAAAAAGAGGCCCTCT 263
145 TTCTCAACAGCAGCAACAGTGCCTGTCATTTCTGTGAGTCTGAAATGTGAAA 204
Db GCCCCCCCCCCCAAGCAAGCAACTATTAAGTCAACAAGTTCACACAAAGTATTA 323
264 GCCCCCCCCCCCAAGCAAGCAACTATTAAGTCAACAAGTTCACACAAAGTATTA 323

Qy 205 CTCTTAAGGATGCTTAAGAACTTGAAACAGTACCGGCAATATTAATCCATATT 264
 Db 324 TTAAATCAACAATATCCAAAGTTATATGAATGATACAGGCTTTAACAAGTCACTTG 383
 Qy 265 GTCCATTCATGAGGCCCATTTTGAAGA 293
 Db 384 GTGCATATCGGTGCGCTTCTTTAAGAGA 412

RESULT 11
 DN689766 1358 bp mRNA linear EST 30-MAR-2005
 LOCUS CGX66-E11.3', mRNA sequence.
 DEFINITION CGX66-E11.3', mRNA sequence.

ACCESSION DN689766
 VERSION DN689766
 KEYWORDS GI:62042478
 SOURCE EST.
 ORGANISM Gasterosteus aculeatus (three spined stickleback)
 Gasterosteus aculeatus
 Gasterosteus aculeatus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphi; Acanthopterygii; Percomorphi; Gasterosteiformes;
 Gasterosteidae; Gasterosteus.

REFERENCE 1 (bases 1 to 1358)
 AUTHORS Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
 Schmutz,J. and Myers,R.M.
 TITLE Expressed sequence tags from Gasterosteus aculeatus
 JOURNAL Unpublished (2003)
 COMMENT Contact: Grimwood, Jane
 Stanford Human Genome Center
 Stanford University School of Medicine
 975 S California Ave, Palo Alto, CA 94304, USA
 Tel: 650 320 5917
 Fax: 650 320 5801
 Email: jane@hgc.stanford.edu

FEATURES
 source
 1..1358
 /organism="Gasterosteus aculeatus"
 /mol_type="mRNA"
 /strain="Comer Creek sticklebacks, WA USA"
 /db_xref="taxon:69293"
 /clone="CGX66-E11"
 /sex="mixed male and female"
 /tissue_type="eyes"
 /dev_stage="adult"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="SHGC-CGX"
 /notes="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTCTAGATCGAGCGGCCGCCCTT25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5' prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxCATC' (where 18 ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at:
 http://www.openbiosystems.com/cdna_library_construction_fa9.php8 The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems:
 http://www.openbiosystems.com/stickleback"

ORIGIN
 Query Match

2.1%; Score 41; DB 8; Length 1358;

Best Local Similarity 49.3%; Pred. No. 4.1; Matches 107; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
 Qy 77 CGCGGTGATGCAAGAGACTTTGATCTGATTTTAAAGAAATTCACACACACCGATG 136
 Db 243 CGGCCCCGGAACAAAGAGTCCATTTTCAATGATGCTTTTCAATTAAGAGGCGC 302
 Qy 137 TCAGTGTTCTTCAAACAGCGACAAACAGGCGCGTGCATTTTCTGTGAGTGGANA 196
 Db 303 CCCCTGCTGCCCCCCCCCCCCCAAGAACTATTAAGTCAAGCGGTACAGTTCACACA 362
 Qy 197 ATGTGAACCTTCTTAAGAGTTGCTAAGAACTTGAACAGTACCGGCAATTAATTAAC 256
 Db 363 GAGTATTTATTCATCAACAATATCCAAAGTGAATGAATGAATAGAGGTTTAACAGTC 422
 Qy 257 TCCATATTCATTCATTCATGAGCCCATTTTGAAGA 293
 Db 423 TACACTGTGATGATCGGTGCGCTTCTTTAAGAGA 459

RESULT 12
 LOCUS CM901981/c
 DEFINITION RPCI42_139E1.TV RPCI-42 Bos taurus genomic clone RPCI42_139E1.
 ACCESSION CM901981
 VERSION CM901981.1 GI:56444251
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 678)
 Larkin,D.M., Donthu,K., Leduc,R., Ryan,K., Liu,L., de Jong,P.J. and
 Lewin,H.A.
 TITLE End sequencing of Holstein BAC library RPCI-42
 JOURNAL Unpublished (2004)
 COMMENT Other GSSs: RPCI42_139E1.TV
 Contact: Harrië Lewin
 Department of Animal Sciences
 University of Illinois at Urbana Champaign
 1201 W. Gregory Dr., Urbana, IL 61801, USA
 Tel: 217 333 5998
 Fax: 217 244 5617
 Email: h-lewin@uiuc.edu
 Clones are derived from the bovine BAC library RPCI-42
 (http://bacpac.chori.org/mbovine42.htm). For BAC library availability, please contact Pieter de Jong
 (pdejong@mail.choi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm).
 Funding was provided in part by grant no. AG2004-34480-14417 from USDA-CSRES (Livestock Genome Sequencing Initiative) and AG58-5438-2-313 from USDA-ARS
 Plate: 139 row: E column: 1
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..678
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /clone="RPCI42_139E1"
 /sex="male"
 /cell_type="blood"
 /clone_lib="RPCI-42"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI-42 Bovine BAC library (Male) produced by Pieter de Jong"

ORIGIN
 Query Match

2.1%; Score 40.8; DB 10; Length 678;

Best Local Similarity 53.8%; Pred. No. 3.8;
Matches 84; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 187 GAGTGGAAATGTGAACCTTTTAAGATTGCTAAAGAACTTAAACAGTACCGCA 246
Db 498 GAGAGAAAAAAGAAAACTCATTAATAAAGATGAGACAACTTCGAAACCTCCGAGAC 439
QY 247 AATAATTAATCTCATTTTGTCCAAATCAATGCGCCCATTTTGAAGAAACAAACAGAG 306
Db 438 AATATTAACGCGCAACACTTCAATCATAGAGTCCCGAAGAGACGAAAGAAAGAC 379
QY 307 CTAGAGACAACTGAGCAAAAGATTCTGACTACATC 342
Db 378 CATGAGAAAAATACTTGAAGAGATTAATAGTTGAAAAC 343

RESULT 13
DR951879 854 bp mRNA linear EST 02-AUG-2005
LOCUS ESR1143418 Aquilegia cDNA library Aquilegia formosa x Aquilegia
DEFINITION pubescens cDNA clone COLRX48, mRNA sequence.
ACCESSION DR951879
VERSION DR951879.1 GI:71721242
KEYWORDS EST.
SOURCE Aquilegia formosa x Aquilegia pubescens
ORGANISM Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Aquilegia.
1 (bases 1 to 854)
Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
Nordborg, M., and Tomkins, J.
Generation of ESTs from Aquilegia
Unpublished (2005)
Other ESTs: ESR1143417
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813
Fax: 805 893 4724
Email: shodges@lifesci.ucsb.edu
Seq primer: M13 Reverse.
FEATURES
SOURCE Location/Qualifiers
1..854
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="COLRX48"
/tissue_type="mixed shoot and floral apical meristems,
flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site 1: SCOR1; Site 2: NotI;
F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
were grown from seed in greenhouses at UC Santa Barbara.
From these plants three sets of tissue were collected: 1)
Small flower buds (<10 mm) and very young inflorescences
(71 & 29% by weight respectively), 2) Medium (7-20 mm) and
large (at or near anthesis) flower buds (65 & 35% by
weight respectively) and 3) Shoot apical meristems. A
fourth set of tissue was collected from plants of A.
formosa. These plants were grown from seed in sand and at
approximately 1 month root tissue and leaf tissue of
various developmental stages were collected (84 & 16% by
weight respectively). Total RNA was extracted from each
set of tissue and pooled in the following proportions:
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
total RNA, mRNA was extracted and enriched for full-length
mRNAs and then normalized with proprietary methods by
Invitrogen."

ORIGIN

Query Match 2.1%; Score 40.4; DB 8; Length 854;
Best Local Similarity 58.2%; Pred. No. 5.4;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 285 TTTAGAGAAACAAACAGAACTAGAGCACTGAGCAAGATTCTGCACTACATCCC 344
Db 403 TTCTGAGAGAAAGAAAGTAAATCAAGAGAAAGATGAAGTAAAGTTAGGATTGACGCA 462
QY 345 TGATATGATATATTTGTCAGTATAGAGGGAGATTTGATCAATAAGTCCGTCATTTGA 404
Db 463 TGATGATGATGATGATGATGAGAGAGATGATGATTTGATTTGATGATTTTACAGAGA 522
QY 405 AC 406
Db 523 TC 524

RESULT 14
AZ145773 426 bp DNA linear GSS 28-AUG-2000
LOCUS AZ145773
DEFINITION SP_0019_B1_H10_SP68 Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library Strongylocentrotus purpuratus
genomic clone Plate=19 Col=19 Row=P, genomic survey sequence.
ACCESSION AZ145773
VERSION AZ145773.1 GI:8297676
KEYWORDS GSS.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidae; Euechinoidae; Echinacea; Echinoidae;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 426)
Cameron, R.A., Mahatras, G., Raat, J.P., Martinez, P., Biondi, T.R.,
Swartzell, S., Wallace, J.C., Pousetka, A.J., Livingston, B.T.,
Wray, G.A., Eftensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H.
A sea urchin genome project: Sequence scan, virtual map, and
additional resources
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
10920195
Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 19 row: P column: 19
Seq primer: SP6
Class: BAC ends
High quality sequence step: 426.
FEATURES
SOURCE Location/Qualifiers
1..426
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone="Plate=19 Col=19 Row=P"
/clone_lib="Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACs; 6; BAC Clones in E-Coli
DH10B"

ORIGIN

Query Match 2.1%; Score 39.8; DB 9; Length 426;
Best Local Similarity 62.6%; Pred. No. 6.5;
Matches 62; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 439 TACAAAATAGATCCCGAGCTTTTCACAAAGCGCATCGAGCTGTGAAAGCGTTGGCG 498
Db 102 TACCTTTTGGTGTGCTATATATTCACAAAGTGACATTCACATCGTTAAACATTTGCA 161
QY 499 CTGATACGAGACGAAACATTAAGAAAGTGCATTAAGA 537

Db 162 CTCGATCCAATTCAGTGCATAAAGAGCTCACAGCAACA 200

RESULT 15

COL16672
10CT16

LOCUS
DEFINITION

ACCESSION
NUMBER

KEYWORDS

SOURCE

ORGANIS

REFERENCE

AUTHORS

TITLE

JOURNAL
COMMENT

2025.03.10

Source:

Query N

Best Local Matchbook

MALCIBIE

공

DB

Qy

공

Search completed: April 8, 2006, 08:09:32
Job time : 6938.08 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:12:00 ; Search time 589.04 Seconds
(without alignments)
5794.037 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920
Sequence: 1 atgagaagaagaagggtgtt.....cgttgcaatgtgaactaa 1920

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6_COMB.seq:*
4: /cgn2_6/prodata/1/ina/7_COMB.seq:*
5: /cgn2_6/prodata/1/ina/8_COMB.seq:*
6: /cgn2_6/prodata/1/ina/9_COMB.seq:*
7: /cgn2_6/prodata/1/ina/10_COMB.seq:*
8: /cgn2_6/prodata/1/ina/11_COMB.seq:*
9: /cgn2_6/prodata/1/ina/12_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1920	100.0	1920	US-09-509-814A-3	Sequence 3, Appl
2	1920	100.0	1920	US-09-920-954-3	Sequence 3, Appl
3	1477.2	76.9	1923	US-09-509-814A-5	Sequence 5, Appl
4	1477.2	76.9	1923	US-09-509-814A-7	Sequence 7, Appl
5	1477.2	76.9	1923	US-09-920-954-5	Sequence 5, Appl
6	1477.2	76.9	1923	US-09-920-954-7	Sequence 7, Appl
7	1136.8	59.2	3003	US-08-873-479-41	Sequence 11, Appl
8	58.2	3.0	7218	US-08-332-463-14	Sequence 14, Appl
9	41.2	2.1	2001	US-09-949-016-2801	Sequence 2801, Ap
10	41	2.1	1141	US-09-806-7088-22	Sequence 22, Appl
11	40	2.1	522	US-09-824-893A-56	Sequence 56, Appl
12	40	2.1	522	US-09-824-893A-70	Sequence 70, Appl
13	40	2.1	1494	US-09-134-000C-2585	Sequence 2585, Ap
14	39.6	2.1	2947	US-09-949-016-815	Sequence 815, Appl
15	39	2.0	522	US-09-824-893A-50	Sequence 50, Appl
16	38.4	2.0	522	US-09-824-893A-37	Sequence 37, Appl
17	38.4	2.0	522	US-09-824-893A-43	Sequence 43, Appl
18	38.4	2.0	522	US-09-824-893A-47	Sequence 47, Appl
19	38.4	2.0	522	US-09-824-893A-60	Sequence 60, Appl
20	38.4	2.0	522	US-09-824-893A-62	Sequence 62, Appl
21	38.4	2.0	522	US-09-824-893A-66	Sequence 66, Appl
22	38.4	2.0	522	US-09-824-893A-76	Sequence 76, Appl
23	38.4	2.0	522	US-09-824-893A-79	Sequence 79, Appl
24	38.4	2.0	522	US-09-824-893A-80	Sequence 80, Appl

25	38.4	2.0	522	US-09-824-893A-83	Sequence 83, Appl
26	38.4	2.0	522	US-09-824-893A-101	Sequence 101, App
27	38.4	2.0	522	US-09-824-893A-112	Sequence 112, App
28	38.4	2.0	522	US-09-824-893A-118	Sequence 118, App
29	38.4	2.0	522	US-09-824-893A-119	Sequence 119, App
30	38.4	2.0	522	US-09-824-893A-120	Sequence 120, App
31	38.4	2.0	522	US-09-824-893A-122	Sequence 122, App
32	38.4	2.0	522	US-09-824-893A-125	Sequence 125, App
33	38.4	2.0	522	US-09-824-893A-130	Sequence 130, App
34	38.4	2.0	522	US-09-824-893A-135	Sequence 135, App
35	38.2	2.0	76810	US-09-949-016-12528	Sequence 12528, A
36	37.2	1.9	399	US-09-621-976-8976	Sequence 8976, Ap
37	36.8	1.9	522	US-09-824-893A-38	Sequence 38, Appl
38	36.8	1.9	522	US-09-824-893A-40	Sequence 40, Appl
39	36.8	1.9	522	US-09-824-893A-46	Sequence 46, Appl
40	36.8	1.9	522	US-09-824-893A-46	Sequence 46, Appl
41	36.8	1.9	522	US-09-824-893A-51	Sequence 51, Appl
42	36.8	1.9	522	US-09-824-893A-52	Sequence 52, Appl
43	36.8	1.9	522	US-09-824-893A-53	Sequence 53, Appl
44	36.8	1.9	522	US-09-824-893A-55	Sequence 55, Appl
45	36.8	1.9	522	US-09-824-893A-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-09-509-814A-3
; Sequence 3, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKATA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUHIKA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
US-09-509-814A-3
Query Match 100.0%; Score 1920; DB 3; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAGAGGTGTTTATCTGTTTATCAGTCGACGATTCGTGACTGT 60
Db 1 ATGAGAAAGAGAGGTGTTTATCTGTTTATCAGTCGACGATTCGTGACTGT 60
QY 61 GCATTAAACATCCCTCGCTGCTGATGACAGAGACTTTGATCTGATTTTAAAGATT 120
Db 61 GCATTAAACATCCCTCGCTGCTGATGACAGAGACTTTGATCTGATTTTAAAGATT 120
QY 121 CAAACAAACAGATGATGATGATTCCTCAACAGCGCAACAAAGATGCGGCTGATTT 180
Db 121 CAAACAAACAGATGATGATGATTCCTCAACAGCGCAACAAAGATGCGGCTGATTT 180

QY 181 CTGGTGAAGTCTGAAAATGTAAGTCTTAAAGATTGCTAAAGAACTTGAACAGTA 240
DB 181 CTGGTGAAGTCTGAAAATGTAAGTCTTAAAGATTGCTAAAGAACTTGAACAGTA 240
QY 241 CCGGCAAAATTAATACTCCATATTGTCTCAATTCAATGGCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAAAATTAATACTCCATATTGTCTCAATTCAATGGCCCATTTTAAAGAAACAAA 300
QY 301 CAGAAAGCTAGAGACAACTGGAGCAAAAGATTCTGACTACATCCCTGATTATGCAATATT 360
DB 301 CAGAAAGCTAGAGACAACTGGAGCAAAAGATTCTGACTACATCCCTGATTATGCAATATT 360
QY 361 GTGCGATATGAGGGGGATGTTGATCAAAAGTCCGCTCCATTTGAACAGCTGAATCAGTG 420
DB 361 GTGCGATATGAGGGGGATGTTGATCAAAAGTCCGCTCCATTTGAACAGCTGAATCAGTG 420
QY 421 GAGCCATACCTGCGCAAAATACAAATAGATCCCAAGCTTTTCACAAAAGGCGCATCGACG 480
DB 421 GAGCCATACCTGCGCAAAATACAAATAGATCCCAAGCTTTTCACAAAAGGCGCATCGACG 480
QY 481 CTGGTGAAGCTTGGCGCTTGTATACGAAAGCAAAATTAAGAAAGTGAATTAAGAGGC 540
DB 481 CTGGTGAAGCTTGGCGCTTGTATACGAAAGCAAAATTAAGAAAGTGAATTAAGAGGC 540
QY 541 ATCGAAGAAATGCTCAAGTACGAGCAAGCAATGACCTCATTTATTTAAGGCAAAAGCT 600
DB 541 ATCGAAGAAATGCTCAAGTACGAGCAAGCAATGACCTCATTTATTTAAGGCAAAAGCT 600
QY 601 GAATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 GAATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCTACGCTTTGTATGAGCAAGCCAGATTTGTCAGATGCGCATGATGATGATGATGATGAT 720
DB 661 AGCTACGCTTTGTATGAGCAAGCCAGATTTGTCAGATGCGCATGATGATGATGATGATGAT 720
QY 721 GGAAGAAACGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GGAAGAAACGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 CTGGGTCGAGCAAGATTAATGCGAATGATGCGAAGCGGTCAATGATGATGATGATGATGAT 840
DB 781 CTGGGTCGAGCAAGATTAATGCGAATGATGCGAAGCGGTCAATGATGATGATGATGATGAT 840
QY 841 GTATTATGAGAAATGCGGCAAGATTAAGAAATGCGAAGCGGTCAATGATGATGATGATGAT 900
DB 841 GTATTATGAGAAATGCGGCAAGATTAAGAAATGCGAAGCGGTCAATGATGATGATGATGAT 900
QY 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 AGCCAAAGCATTCAGTGAAGTGCAGAAATTCATCAAACTCTGAGGGGCGAGCGGTGAT 1020
DB 961 AGCCAAAGCATTCAGTGAAGTGCAGAAATTCATCAAACTCTGAGGGGCGAGCGGTGAT 1020
QY 1021 GGGGCTTACAGCAAGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GGGGCTTACAGCAAGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ATTTCTTTTCGCGGCTGGGAATGAAGAGCGGAGGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 ATTTCTTTTCGCGGCTGGGAATGAAGAGCGGAGGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GCTAATAAAAGCCATTAACAGTGGCGCAACCGAAACCTGCGTCAAGCTTGGTTCCTAT 1200
DB 1141 GCTAATAAAAGCCATTAACAGTGGCGCAACCGAAACCTGCGTCAAGCTTGGTTCCTAT 1200
QY 1201 GCAAGTATATTAACAGTGGTCAAGTTCCTTCCGTTGGGCGCAAAAGATGGGCGA 1260
DB 1201 GCAAGTATATTAACAGTGGTCAAGTTCCTTCCGTTGGGCGCAAAAGATGGGCGA 1260

QY 1261 ATCAAGCTGATGTCATGAGCGCAGGAGCATACATTTTATCAGCAAGATCTTCTTCA 1320
DB 1261 ATCAAGCTGATGTCATGAGCGCAGGAGCATACATTTTATCAGCAAGATCTTCTTCA 1320
QY 1321 CCGGATTCCTCTCTTGGGCGAATCATGACAGCAAAATATGCTTATATGAGTGAACGTCC 1380
DB 1321 CCGGATTCCTCTCTTGGGCGAATCATGACAGCAAAATATGCTTATATGAGTGAACGTCC 1380
QY 1381 ATGCGACACCGAATTTGTTGGCGGGAATGTTGCAAGCTCCGTTAGCATTTTGTAAAAAT 1440
DB 1381 ATGCGACACCGAATTTGTTGGCGGGAATGTTGCAAGCTCCGTTAGCATTTTGTAAAAAT 1440
QY 1441 AGAGGAATCATCTCTAAGCTTCCCTATTTGAAGACGCTTGAATGACAGTGTCTGAT 1500
DB 1441 AGAGGAATCATCTCTAAGCTTCCCTATTTGAAGACGCTTGAATGACAGTGTCTGAT 1500
QY 1501 GTTGAATTTGGTATTCGAAACGAAACCAAGATGGGCGAGTGAACCTGATTAATTCG 1560
DB 1501 GTTGAATTTGGTATTCGAAACGAAACCAAGATGGGCGAGTGAACCTGATTAATTCG 1560
QY 1561 TTGAACGTTGCTATGTAAGCAAGATCCAGTCCCTATCACTAGCCAAAGGCAATAT 1620
DB 1561 TTGAACGTTGCTATGTAAGCAAGATCCAGTCCCTATCACTAGCCAAAGGCAATAT 1620
QY 1621 ACCTTACTGCAAGCGCGGCAAGCATTTGAAATCTCCCTGATGAGTGGAGTCCCT 1680
DB 1621 ACCTTACTGCAAGCGCGGCAAGCATTTGAAATCTCCCTGATGAGTGGAGTCCCT 1680
QY 1681 GCAAGCACTACTGCTTCTGTAACCTGCTCAATGATTTGATTTGTCATTAACAGCA 1740
DB 1681 GCAAGCACTACTGCTTCTGTAACCTGCTCAATGATTTGATTTGTCATTAACAGCA 1740
QY 1741 AACGGAACAAGATATGTCGGGAATGATCTTCACACACATTTGACATTAATGAGGAG 1800
DB 1741 AACGGAACAAGATATGTCGGGAATGATCTTCACACACATTTGACATTAATGAGGAG 1800
QY 1801 CGCATATAAGTGAAGAAATGATTTATTAATTCGCCCCCAAGTGAACATATACATTGAG 1860
DB 1801 CGCATATAAGTGAAGAAATGATTTATTAATTCGCCCCCAAGTGAACATATACATTGAG 1860
QY 1861 GTGCAAGCATTAATGTCGGGTTGACCACAAAATCTTCTGTTGCAATTTGTAATTA 1920
DB 1861 GTGCAAGCATTAATGTCGGGTTGACCACAAAATCTTCTGTTGCAATTTGTAATTA 1920

RESULT 2
US-09-920-954-3
Sequence 3, Application US/09920954
Patent No. 6759228
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT FILING DATE: US/09/920, 954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA

ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1).. (1920)
US-09-920-954-3

Query Match 100.0%; Score 1920; DB 3; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGAAAGAAAGAGTGTATCTGTTTATCGAGTGCAGAGTCTGTGCACTGTT 60
DB 1 ATGAGAAAGAAAGAGTGTATCTGTTTATCGAGTGCAGAGTCTGTGCACTGTT 60
QY 61 GCATTAAACATCCCTCGGTGTGTGCAAGACCTTTGATCTGATTTTAAAGAAAT 120
DB 61 GCATTAAACATCCCTCGGTGTGTGCAAGACCTTTGATCTGATTTTAAAGAAAT 120
QY 121 CAACCAACACGATGTCAGTGGTTCTCCAAACAGGCAAAACAGGTGCGCTGATTT 180
DB 121 CAACCAACACGATGTCAGTGGTTCTCCAAACAGGCAAAACAGGTGCGCTGATTT 180
QY 181 CTGTGTAGTCTGAAAAATGTGAACCTTTAAAGATTGCTAAAGAACTTGAAACATA 240
DB 181 CTGTGTAGTCTGAAAAATGTGAACCTTTAAAGATTGCTAAAGAACTTGAAACATA 240
QY 241 CCGGCAAAATTAATATCCATATTTGTCATTCATATGCGCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAAAATTAATATCCATATTTGTCATTCATATGCGCCCATTTTAAAGAAACAAA 300
QY 301 CAGAGCTAGAGCAACCTGAGCAAGATTTCTGACATACCTCCGATTAATGATATATT 360
DB 301 CAGAGCTAGAGCAACCTGAGCAAGATTTCTGACATACCTCCGATTAATGATATATT 360
QY 361 GTCGATATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCATTTGAAACATGAGATCAGT 420
DB 361 GTCGATATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCATTTGAAACATGAGATCAGT 420
QY 421 GAGCCATCTTGGCCGAATTAACAAATATGATCCCACTTTTCACAAAAGGCGATCAGC 480
DB 421 GAGCCATCTTGGCCGAATTAACAAATATGATCCCACTTTTCACAAAAGGCGATCAGC 480
QY 481 CTGGTGAAGCGTGGCGCTTGATACGAAGCAGAAATTAAGCAATTAAGAGC 540
DB 481 CTGGTGAAGCGTGGCGCTTGATACGAAGCAGAAATTAAGCAATTAAGAGC 540
QY 541 ATCGAGGAATCGCTCAGTACGTAGCAACATGACGTCCATTAATTAACGCAAGCCT 600
DB 541 ATCGAGGAATCGCTCAGTACGTAGCAACATGACGTCCATTAATTAACGCAAGCCT 600
QY 601 GAATTAATAGGTATGATATGATGTGGCAGAGGTAATGTCAAGCGGATGTGGCAGAGC 660
DB 601 GAATTAATAGGTATGATATGATGTGGCAGAGGTAATGTCAAGCGGATGTGGCAGAGC 660
QY 661 AGCTACGGTTTGTATGAGCAAGGCGAGATTTGTGCGAGTTGCCGATCTGAGATTA 720
DB 661 AGCTACGGTTTGTATGAGCAAGGCGAGATTTGTGCGAGTTGCCGATCTGAGATTA 720
QY 721 GGAAGAAACGACGATTCAGTATGATGAGGCTTCCGCGGTAAATTAACAGCATATATGCA 780
DB 721 GGAAGAAACGACGATTCAGTATGATGAGGCTTCCGCGGTAAATTAACAGCATATATGCA 780
QY 781 CTGGGTCCGACGAATATATGCGAATGACGATCATGTGATACCCATGTGGCAGGTTGG 840
DB 781 CTGGGTCCGACGAATATATGCGAATGACGATCATGTGATACCCATGTGGCAGGTTGG 840
QY 841 GATATTGAGAAATGGGCGCAAGATTAAGGATGCGACCTCAAGCGGATCTGTTTTTCA 900
DB 841 GATATTGAGAAATGGGCGCAAGATTAAGGATGCGACCTCAAGCGGATCTGTTTTTCA 900
QY 901 TCCATCATATGATAGCAGTGTGGGTCTTGAGGCTTGCTTCCATCTGCAACCTTAATTC 960
DB 901 TCCATCATATGATAGCAGTGTGGGTCTTGAGGCTTGCTTCCATCTGCAACCTTAATTC 960
```

```
QY 961 AGCCAAAGCATTCAGTGCAGGTGCAGAAATTCATCAAACTCTGCGGGGCGACGGGTGAAT 1020
DB 961 AGCCAAAGCATTCAGTGCAGGTGCAGAAATTCATCAAACTCTGCGGGGCGACGGGTGAAT 1020
QY 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATATATGAAAAATGATATAGACG 1080
DB 1021 GGGGCTTACACGACAGATTCAGAAATGTGATGATATATGAAAAATGATATAGACG 1080
QY 1081 ATTCCTTTTGGCGGTGGGAATGAAAGGCGGAAACGGGATACATCAGTGCACCTGTGACG 1140
DB 1081 ATTCCTTTTGGCGGTGGGAATGAAAGGCGGAAACGGGATACATCAGTGCACCTGTGACG 1140
QY 1141 GCTAAAGACCCATTAACATCGCGGCAACCGAAACCTCGGTCCAAAGCTTGGTCTTAT 1200
DB 1141 GCTAAAGACCCATTAACATCGCGGCAACCGAAACCTCGGTCCAAAGCTTGGTCTTAT 1200
QY 1201 GCGATTAATTAATTAACGATGTCAGTGTCTCTCCGTTGGGCGGCAAAAGATGGGCGA 1260
DB 1201 GCGATTAATTAATTAACGATGTCAGTGTCTCTCCGTTGGGCGGCAAAAGATGGGCGA 1260
QY 1261 ATCAAGCTGATATGATCAATGCGCGCAGGGAATTAATTAATGACAAAGATCTTCTTGA 1320
DB 1261 ATCAAGCTGATATGATCAATGCGCGCAGGGAATTAATTAATGACAAAGATCTTCTTGA 1320
QY 1321 CCGAATCTCTCTTCTGGGCGAATCAATGACAGCAATATGCTTAATGAGTGAACGTCC 1380
DB 1321 CCGAATCTCTCTTCTGGGCGAATCAATGACAGCAATATGCTTAATGAGTGAACGTCC 1380
QY 1381 ATGGCAACGCGATTTGGCGGGGAAATGTTGACAGCTCCGATGAGATTTTGAAGAAAT 1440
DB 1381 ATGGCAACGCGATTTGGCGGGGAAATGTTGACAGCTCCGATGAGATTTTGAAGAAAT 1440
QY 1441 AGAGGAATCACTCTTAAGCTTCCCTAATTAAGAGAGCTTTGATTTGACAGTCTGTGAT 1500
DB 1441 AGAGGAATCACTCTTAAGCTTCCCTAATTAAGAGAGCTTTGATTTGACAGTCTGTGAT 1500
QY 1501 GTTGAATGGGTTATCCGAACGAAACCAAGGATGGGCGGAGTGAACCTTGATTAATCG 1560
DB 1501 GTTGAATGGGTTATCCGAACGAAACCAAGGATGGGCGGAGTGAACCTTGATTAATCG 1560
QY 1561 TTGAACGTTTCCCTAATGTAACGAATCCAGTCCCTATCAATGACCAAAACGACATAT 1620
DB 1561 TTGAACGTTTCCCTAATGTAACGAATCCAGTCCCTATCAATGACCAAAACGACATAT 1620
QY 1621 ACCCTTAATGCAACGGCGGCAAGCATTGAATACTCCCTGTATGTGCGAGTCCCT 1680
DB 1621 ACCCTTAATGCAACGGCGGCAAGCATTGAATACTCCCTGTATGTGCGAGTCCCT 1680
QY 1681 GCAAGCACTACTGCTTCTGTAACTCCTGTGCAATGATTTGATTTGATCAAGACCA 1740
DB 1681 GCAAGCACTACTGCTTCTGTAACTCCTGTGCAATGATTTGATTTGATCAAGACCA 1740
QY 1741 AACGGAACGAATATGTGGGAAATGACTTCTAGCAACATTTGAATTAATCTGGGATGCG 1800
DB 1741 AACGGAACGAATATGTGGGAAATGACTTCTAGCAACATTTGAATTAATCTGGGATGCG 1800
QY 1801 CGCAATTAACGTGAATAATGATTTTAATTAATGCCCCCAAGGGAACATATACATTGAG 1860
DB 1801 CGCAATTAACGTGAATAATGATTTTAATTAATGCCCCCAAGGGAACATATACATTGAG 1860
QY 1861 GTGCAAGCATATATATGTCGGGTGAGCACAAAATCTCGTTGGCAATTTGAACTTA 1920
DB 1861 GTGCAAGCATATATATGTCGGGTGAGCACAAAATCTCGTTGGCAATTTGAACTTA 1920
```

RESULT 3
US-09-509-814A-5
Sequence 5, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
OKUDA, MITSUYOSHI

Db 1750 ACGACATATGTAGAAATGACTTACTTCCGCATACATGATTAACGTGGATGCCCAT 1809
Qy 1807 AACGTGAAATATATTTATTAATTCGCCCAAGTGAAATATACATTAAGGTGCA 1866
Db 1810 AACGTGAAATATATTTATTAATTCGCCCAAGTGAAATATTAAGGTGCA 1869
Qy 1867 GCATATAATGTCCGGTGGACCAAACTTCTCGTGGCAATTTGTGAATTA 1920
Db 1870 GCTTAATGATGACGGGTGGACCAAGACTTCTCGTGGCAATTTGAATTA 1923
RESULT 4
US-09-509-814A-7
Sequence 7, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OR INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent in version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1923)
US-09-509-814A-7
Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
Qy 7 AAGAGAGAGGTGTTTATCTGTTTATGAGCTGACGCAATTCGTGCACTGTTCATTA 66
Db 10 AAGAAAGAGGTGTTTATCTGTTTATGAGCTGACGCAATTCGTGCACTGTTCATTA 69
Qy 67 AACATCCCTCGGCTGCTGATGCAAGCACTTTGATCTGATTTTAAAGAAATCAACA 126
Db 70 AGTATTCATCTGCGAGGTGCTGCAAGCAATTTGATCTGATTTTAAAGAAATCAACA 129
Qy 127 ACAACCGATGTCAGTGTGTTCTTCCAAACGCAAAACAGTGTGCGGTGCTGCTG 186
Db 130 ACAACCTGATGCTAAAGTGTCTTCCAAAGCGGGCAGACTGTGTGCTGCTTCTG 189
Qy 187 GAGTCTGAATAATGTAACTTCTTAAGGATTTGCTAAAGAACTTGAACATACCGGA 246
Db 190 GAATCTGAAATGTAAACTCCCAAAAGTGTGCAAGAAAGCTTGAACAGTCCCGCA 249
Qy 247 AATAATTAATCTCATATTTGTCATTCATTCATTCATTCATTCATTCATTCATTCAT 306
Db 250 AATAATTAATCTCATATTTGTCATTCATTCATTCATTCATTCATTCATTCATTCAT 309
Qy 307 CTAGAGCAACTGAGCAAAAGATTTCTGACATTCATTCATTCATTCATTCATTCATTC 366
Db 310 CTGGAAGAAACAGGGGCAAAAGATTTCTGACATTCATTCATTCATTCATTCATTCAT 369
Qy 367 TATGAGGGGAGATTTCAATCAAAAGTCCGCTCATTTGAACAGTGAATCACTGAGCA 426

Db 370 TATGAGGGGAGATTTCAATCAAAAGTCCGCTCATTTGAACAGTGAATCCGAGCTT 429
Qy 427 TACTTCCCAATATCAAAATAGATCCCAAGTCTTTCACAAAGGCCCATGACGCTGTG 486
Db 430 TATTTGCCGATATACGAATATATGCCAGCTTTTCACAAAGGGGCATCAAGCTTTGTA 489
Qy 487 AAAGCTTGGCGCTGTATACGAGCAAGAACATTAAGAAATGCAATTAAGAGCATGAG 546
Db 490 AAAGCAGTGGCGCTGTATACGAGCAAGAAATTAAGAGTGCATTAAGAGCATGAG 549
Qy 547 GAAATGCTCAGTACGTACGAGCAATGACGTCCATTAATTAACGCAAGCTGAAATAT 606
Db 550 CAATTCGACAAATTCGCAATTAAGCAATGATGCTATATTAATTCGCAAGCTGATAT 609
Qy 607 AAGGTGATGAATGATGTGCGCAGAGTATTTCAAGCGAATGTGCAAGAGCAAGCTAC 666
Db 610 AAGGTGATGAATGATGTGCGCAGAGTATTTCAAGCGAATGTGCTCAGAGCAAGCTAC 669
Qy 667 GGTGTGTATGACAAAGCCAGATTTGTGCAATTTGCCGATATCTGATTTGANTACAG 726
Db 670 GGTGTGTATGACAAAGCAAGATCTGAGCGTTCGATACAGGCTGTATACAGGTCCG 729
Qy 727 AACGACAGTTGATGATGATGAAGCTTCCGCGGTAAATTAACGACATATATGACTGG 786
Db 730 AATGACAGTTGATGATGATGAAGCTTCCGCGGAAATTAATGCAATTAATGATGAG 789
Qy 787 CGGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
Db 790 CGGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 849
Qy 847 GGAATATGCGGCAAGCAATTAAGAAATGCACTTCAAGCCGAATCTGATTTTCAATCAT 906
Db 850 GGAATATGCGGCAAGCAATTAAGAAATGCACTTCAAGCCGAATCTGATTTTCAATCAT 909
Qy 907 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Db 910 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
Qy 967 GCATTCAGTGCAGTGTGCGCAATTCATCAAACTCTGCGGGGCGAGCGGTGATGGGG 1026
Db 970 GCATTCAGTGTGCGCAATTCATCAAACTCTGCGGGGCGAGCGGTGATGGGGCT 1029
Qy 1027 TACACGACGATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1086
Db 1030 TACACGACGATTCGAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1089
Qy 1087 TTGCGGCTGCGGAAATGAAGAGCGCGAAGCGGTGATCAATCAAGTGCATCGCTAA 1146
Db 1090 TTGCGTGCCTGGAAATGAAGAGCGCGAAGCGGTGATCAATCAAGTGCATCGCTAA 1149
Qy 1147 AACGCAATTAACAGTGTGCGCAACGAAACCTGCTCAAGCTTGGTCTTATGCAAT 1206
Db 1150 AATGCAATTAACAGTGTGCGCAACGAAACCTGCTCAAGCTTGGTCTTATGCAAT 1209
Qy 1207 AATATTAACAGTGTGCAAGTCTTCCGCTGCGCGCAAAAGATGGGAAATCAAG 1266
Db 1210 AATATCAACAGTGTGCAAGTCTTCCGCTGCGCGCAAAAGATGGGAAATCAAA 1269
Qy 1267 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
Db 1270 CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1329
Qy 1327 TCTTCTCTTCTGCGGCAATTCATGACAGCAATATGCTTATATGAGTGAAGCTGATG 1386
Db 1330 TCTTCTCTTCTGCGGCAATTCATGACAGTAAATATGATATCAATGAGTGAAGCTGATG 1389
Qy 1387 AACCGAATTTGTGCGGGAATGTGACAGCTCCGCTGAGCAATTTGTGAATAATGAG 1446
Db 1390 AACCGAATTTGTGCGGGAATGTGACAGCTCCGCTGAGCAATTTGTGAATAATGAG 1449
Qy 1447 ATCACTCTAAGCCCTTCCCTATTAAGAACAGCTTTGATGAGAGTGTGATGATGAT 1506

Db 1450 ATCAACCAAGGCTTCTCTATTAAAGGGGCACTGATTTGCCGGTGCAGCTGACATCGGC 1509
Qy 1507 TTGGGTATCCGAAACGAAACCAAGAGTGGGGCCGAGTGAACCTGGATTAATCGTTGAAC 1566
Db 1510 CTGGCTACCCGAAACGATACCAAGAGTGGGAGCAGATGACATTTGAATTAATCCCTGAAC 1569
Qy 1567 GTTGCCTATGTAAGAAAGATCCAGTGGCCCTATCACTAGCCAAAGGCAATATACCTTT 1626
Db 1570 GTTGCCTATGTAAGAAAGATCCAGTGGCCCTATCACTAGCCAAAGGCAATATACCTTT 1629
Qy 1627 ACTGCAACGCGCGGCAAGCCATTTGAAATCTCCGTATGATGGATGCCCTCCGCAAGC 1686
Db 1630 ACTGCTATCGCCGCAAGCTTTGAAATCTCCGTATGATGGATGCCCTCCGCAAGC 1689
Qy 1687 ACTACTGCTTCTGTAACCTGCTCAATGATTTGGATTGGTCAATTAACGACCAACGGA 1746
Db 1690 ACAACTGCTTCCGTAACGCTTGTCAATGATCTGGACCTTGTCAATTAACGCTCAATGAC 1749
Qy 1747 ACAAGATATGTCGGGAATGACTTCCAGACCAATTTGCAATTAACCTGGGATGGCCGCAAT 1806
Db 1750 ACAAGATATGTAAGAAATGACTTACCTTCCATACATGATTAACCTGGGATGGCCGCAAT 1809
Qy 1807 AACGTAGAAATGATTTTATTAATTCGCCCAAGTGAACATATATCAATGAGGTGCA 1866
Db 1810 AACGTAGAAATGATTTTATTAATGACCAAGGCGGAGGTATTAATTAATGAAATGACAG 1869
Qy 1867 GCATTAATGTCGCCGTTGGACCAAAATCTTCTGTTGGCAATTTGACTTA 1920
Db 1870 GCATTAATGTCGCCGTTGGACCAAAATCTTCTGTTGGCAATTTGACTTA 1923

RESULT 5

US-09-920-954-5
Sequence 5, Application US/09920954
Patent No. 6759228

GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-5

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAAAGAGGTGTTTATCTGTTTATCAGTGCAGAGGATTTCTGACCTGTTGATTA 66
Db 10 AAGAAAGAGGTGTTTATCTGTTTATCAGTGCAGAGGATTTCTGACCTGTTGATTA 69

Qy 67 AACATCCCTGGCTGATGATGCAAGGACTTTTGAATCTGATTTTAAAGAAATTCAAACA 126
Db 70 AGTAATCATCTGACGGGTGGCAAGGATTTTGAATCTGATTTTAAAGAAATTCAGACA 129
Qy 127 AACACCAATGTCAGTGTGTTCTCCAAACAGGCAAAACAGGTGGGCTGATTTCTGGTG 186
Db 130 AACACTGATGTAAGGTTTCTCCAAAGAGGGGAGAGCTGGTGTCTGCTTTCTGGTG 189
Qy 187 GAGTCTGAATGTAAGACTTCTTAAGAGATTGCTTAAGAACTTGAACAGTACGCGCA 246
Db 190 GAATCTGAATGTAAGACTTCCAAAGGTTTGAAGAAAGTGAACAGTCCCGCA 249
Qy 247 AATAATAATCCCATATTTGTCATTAATGCAATGAGCCCATTTTGAAGAAACAGAG 306
Db 250 AATAATAATCCCATATTTATCAATTAATCAATGACCAATTTTGAAGAAACAGAG 309
Qy 307 CTAGAGCAACTGAGCAAAAGATTCTGACTACATCCCTGATTAATATATTTGTCAG 366
Db 310 CTGAAAGAAACAGGGGCAAAAGATTCTGACTACATCCTGATTAATGCTTACATTTGTCAG 369
Qy 367 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCATGAGCCA 426
Db 370 TATGAGGGGAGTGTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCCGTGAGCCT 429
Qy 427 TACTTGGCGAAATCAAAATATGATCCCGAGCTTTCAAAAGGCGCATGACGCTGGTG 486
Db 430 TATTGGCGAATACAGATATGATCCCGAGCTTTCAAAAGGCGCATGAGCTTTGA 489
Qy 487 AAACGTTGGCGCTTGTATACAGACAGAACATTAAGAGTGAATTAAGAGGATCGAG 546
Db 490 AAACAGTGGCGCTTGTATACAGACAGAAATTAAGAGGTTGCAATTAAGAGGATCGAA 549
Qy 547 GAATTCGCTCAGTACGTAACAGCAAGTGAAGTCAATTAATTAAGCAAAAGCTGAAAT 606
Db 550 CAATTCGCAATTCGCAATTAAGCAATGATGTGTATTAATTAAGCAAAAGCTGAAAT 609
Qy 607 AAGGTATGAATGATGTCGCGAGAGTATTGTCAAAGCGGATGTCAGACAGAGCTAC 666
Db 610 AAGGTATGAATGATGTCGCGAGAGTATTGTCAAAGCGGATGTCAGAGCTAC 669
Qy 667 GGTTTGATGCAAGGCGAGATTTGTGCAAGTTGCGCATTACTGATTTGATACAGAGA 726
Db 670 GGTTTGATGCAAGGCGAGATTTGTGCAAGTTGCGCATTACTGATTTGATACAGAGT 729
Qy 727 AAGCAAGTTGATGCAATGAAGCTTTCGCGGTAAATTAACAGCACTATATGCTGGGT 786
Db 730 AATGACAGTTGATGCAATGAAGCTTTCGCGGTAAATTAACAGCACTATATGCTGGGA 789
Qy 787 CGACGAATTAATGGAATGATACAGACGTCATGTCATCCCATGTCGAGGTTCCGTAATTA 846
Db 790 CGACGAATTAATGGAATGATACAGACGTCATGTCATCCCATGTCGAGGTTCCGTAATTA 849
Qy 847 GAAATGCGCGCAAGATTAAGAAATGCGACCTCAAGCGAATCTGTTTTCAATCCATC 906
Db 850 GAAATGCGCGCTCACTAATTAAGAAATGCGACCTCAAGCGAATCTGTTTTCAATCCATC 909
Qy 907 ATGATATGCAATGATGAGCTTGGAGCTTGCCTTCAATCTGCAAACTTATTAAGCCAA 966
Db 910 ATGATATGCAATGATGAGCTTGGAGCTTGCCTTCAATCTGCAAACTTATTAAGCCAA 969
Qy 967 GCATTCAGTGCAGGTGCGAATTCATACAAACTCTGGGGGCGAGCGGTGAATGGGGCC 1026
Db 970 GCATTCAGTGCAGGTGCGAATTCATACAAACTCTGGGGGCGAGCGGTGAATGGGGCC 1029
Qy 1027 TACACGACAGATTCAGAAATGTCATATGATTAAGAAATTAATGATGATGATCTT 1086
Db 1030 TACACGACAGATTCAGAAATGTCATATGATTAAGAAATTAATGATGATGATCTT 1089
Qy 1087 TTGCGCGCTGGGATTAAGGCGCGACAGCGGTACCATTCAGTGAACCTGATACGCTTAA 1146
Db 1090 TTGCGCTCCGGGAATGAAGACCGAAGCGGGAACCATTCAGTGAACCGGCTTAA 1149
Qy 1147 AACGCATTAACAGTGGCGCAACGAAACCTGTGCTCAAGCTTCCGTTCTATGAGAT 1206

Db 1150 AATGCAATATACAGTCGAGCTACGAGAAACCTCCGCCAAGCTTGGCTTATCGGAC 1209
Qy 1207 AATATTAACCACTGTCACAGTTCTCTTCCGTCGCCGCAAAAGATGGCGAATCAAG 1266
Db 1210 AATATCAACCATGTGGCAAGTTCTCTTCACTGAGCCGCAAAAGATGAGCGATCAAA 1269
Qy 1267 CCTGATGTCAATGGCGCAGGAGCATATTTATTCAGCAAGATCTTCTTTCGACCCGAT 1326
Db 1270 CCGAATGTCAATGGCAACGGGAGAGTCAATACAGCAAAATCTTCTTTCGACCCGAT 1329
Qy 1327 TCCTCTCTTGGCGCATCAATGACAGCAATATGCTATATGGGTGGAAAGTCATAGCA 1386
Db 1330 TCCTCTCTTGGCGCAACATGACAGTAATATGATGATGATGATGATGATGATGATG 1389
Qy 1387 ACACCGATGTGGCGGAGAAATGTTGCAACAGCTCCGTCGAGCAATTTTGAATAAGCA 1446
Db 1390 ACACCGATGTGGCTGAGAAAGTGGCAACAGCTTCGTCGAGCAATTTTGAATAAGCA 1449
Qy 1447 ATGACTCTTAAGCTTCCCTATGAAAGCAAGCTTGTATGTCAGAGTCTGTGATGTTGA 1506
Db 1450 ATGACACCAAAAGCTTCTCTATTAAGCGGCACTGATGCGGTGACAGTCAATCGGC 1509
Qy 1507 TTGGCTTATCCGAAACGAAACCAAGAGATGGGCGGAGTGAACCTGATTAATGTTGAAC 1566
Db 1510 CTGGCTTACCCGAAACGATACCAAGAGATGGGAGAGATGATGATGATGATGATGATG 1569
Qy 1567 GTTGCCTATGTGAACCAATCAGATGCTTATCACTGATGATGATGATGATGATGATG 1626
Db 1570 GTTGCCTATGTGAACCAATCAGATGCTTATCACTGATGATGATGATGATGATGATG 1629
Qy 1627 ACTGCAACGCGCGGCAAGCAATGAAATCTCCCTGATGATGATGATGATGATGATGATG 1686
Db 1630 ACTGCAACGCGCGGCAAGCAATGAAATCTCCCTGATGATGATGATGATGATGATGATG 1689
Qy 1687 ACTACTGCTTCTGTAACCTGTCATGATGATGATGATGATGATGATGATGATGATGATG 1746
Db 1690 ACAACGCTTCCGTAACGCTGTCATGATGATGATGATGATGATGATGATGATGATGATG 1749
Qy 1747 ACAAGATGTGCGGGAATGATCTTCTGAGCAACATTTGACATTAAGTGGAGTGGCAG 1806
Db 1750 ACAAGATGTGCGGGAATGATCTTCTGAGCAACATTTGACATTAAGTGGAGTGGCAG 1809
Qy 1807 AACGTGAAGAAATGATTTATTAATGCGCCCAAGGGAACATTAATCAATGAGTGGCA 1866
Db 1810 AACGTGAAGAAATGATTTATTAATGCAACAAAGGGAACGATTAATCAATGAGTGGCA 1869
Qy 1867 GCATTAATGTGCGGTTGGAACCAAACTTCTGTTGGCAATTTGAACTTA 1920
Db 1870 GCTTATTAAGTACGAGTGGAGCAACAGCTTCTGTTGGCAATTTGAACTTA 1923

RESULT 6
US-09-920-954-7
Sequence 7, Application US/09920954
Patent No. 6759228
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIKA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGIYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-7

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAAAGAGTGTATCTGTTTATGAGTGGAGGATCTGTGCACTGTGATTA 66
Db 10 AAGAAAGAGTGTATCTGTTTATGAGTGGAGGATCTGTGCACTGTGATTA 69
Qy 67 AACATCCCTCGAGTGTGATGCAAGCACTTTGATCTGATTTTAAAGAAATCAACA 126
Db 70 AGTAATCCATCTGCAAGTGTGCAAGGATTTGATCTGATTTTAAAGAAATCAACA 129
Qy 127 ACACCGATGTGATGTTTCTTCAACAGCAACAAAGTGGCTGATTTCTGTG 186
Db 130 ACACCGATGTGATGTTTCTTCAACAGGAGGAGATGATGATGATGATGATGATG 189
Qy 187 GAGTCGAAAGATGAACTTCTTAAGGATGCTTAAGAACTTGAACAGTACCGGA 246
Db 190 GATTCGAAAGATGAACTTCTTAAGGATGCTTAAGAACTTGAACAGTACCGGA 249
Qy 247 AATTAATTAATCTCATATGTCATATTCATATGAGCCCATTTTGAAGAAACAAAG 306
Db 250 AATTAATTAATCTCATATATTCATATTCATATGAGCAATTTTGAAGAAACAAAG 309
Qy 307 CTAGAGCAATCTGAGCAAGAAATTTCTGACATATCTGATTAATGATTAATGTCAG 366
Db 310 CTAGAGCAATCTGAGCAAGAAATTTCTGACATATCTGATTAATGATTAATGTCAG 369
Qy 367 TATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 426
Db 370 TATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 429
Qy 427 TACTTCCGAAATTAACAAATAGATCCCGACTTTTCACAAAGGCGATCGAGTGTG 486
Db 430 TATTTCCGATATACGAATAGATCCCGACTTTTCACAAAGGCGATCGAGTGTG 489
Qy 487 AAGCGTGGCGCTTGTATGAGAGCAACAAATTAAGAGTGCATTAAGAGCATCGAG 546
Db 490 AAGCGTGGCGCTTGTATGAGAGCAACAAATTAAGAGTGCATTAAGAGCATCGAG 549
Qy 547 GAATTCCTCACTAGTACAGCAAGCAATGATGATGATGATGATGATGATGATGATG 606
Db 550 GAATTCCTCACTAGTACAGCAAGCAATGATGATGATGATGATGATGATGATGATG 609
Qy 607 AAGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 666
Db 610 AAGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 669
Qy 667 GGTGTGATGAGCAAGCGCAATGTCGAGTGGCGATCTGATGATGATGATGATGATG 726
Db 670 GGTGTGATGAGCAAGCGCAATGTCGAGTGGCGATCTGATGATGATGATGATGATG 729
Qy 727 AACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
Db 730 AATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 789
Qy 787 CGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
Db 790 CGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 849


```

QY 847 GGAATGCGGCAAGAAATGAAGAAATGCGACCTCAAGCGAATCTGGTTTTCATCCATC 906
DB 850 GGAATGCGGCTCCATTAATTAAGAAATGCGCTCAAGCGAATCTGGTTTTCATCCATC 909
QY 907 ATGATAGCAATGATGCGGCTTGAAGGCTTGCCTCCATCTGCAAACTTTATTCAGCCA 966
DB 910 ATGATAGCGATGCGGGAATTTGAAGGACTACCTTGAATCTGCAAACTTTATTCAGCCA 969
QY 967 GCATTCAGTGCAGTGGCCGAAATTCATCAAACTCTGGGGGGCGAGCGGTGAATGGGCT 1026
DB 970 GCATTCAGTGCAGTGGCCGAAATTCATCAAACTCTGGGGGGCGAGCGGTGAATGGGCT 1029
QY 1027 TACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
DB 1030 TACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
QY 1087 TTCCGCGGCTGGAAATGAAGCCGAAACGCGGTACCATCATGTCACCTTGTACGCTAAA 1146
DB 1090 TTCCGCGGCGGAAATGAAGCCGAAACGCGGTACCATCATGTCACCTTGTACGCTAAA 1149
QY 1147 AACGCAATACAGTGGCGGCAACCGAAACCTGCGCTCCAGCTTGGTTCTATGCAAT 1206
DB 1150 AATGCAATACAGTGGCGGCAACCGAAACCTGCGCTCCAGCTTGGTTCTATGCAAT 1209
QY 1207 AATATTAACAGTGGCAAGTTCTTCCGCGGCGGCAAAAGATGGCGGAATCAAG 1266
DB 1210 AATATTAACAGTGGCAAGTTCTTCCGCGGCGGCAAAAGATGGCGGAATCAAG 1269
QY 1267 CCTGATGTCATGGCGCCGAGGACATACATTTTATGACAAAGATCTTCTTTCAGCCGAT 1326
DB 1270 CCTGATGTCATGGCGCCGAGGACATACATTTTATGACAAAGATCTTCTTTCAGCCGAT 1329
QY 1327 TCTCTCTTCTGGGCGCAATCAACAGCAAAATGCTATATGGGTGAACCTTCATGCA 1386
DB 1330 TCTCTCTTCTGGGCGCAACATCAAGTAATATGATACATGGGTGAACCTTCATGCT 1389
QY 1387 ACACGATTTGTCGCGGGAATGTCACAGCTCCGCGGAGCAATTTGTGAAAAATGAGAA 1446
DB 1390 ACACGATTTGTCGCGGGAATGTCACAGCTCCGCGGAGCAATTTGTGAAAAATGAGAA 1449
QY 1447 ATCACTCTTAAGCCTTCCCTATTAAGCAAGCTTGTGATGATGCTGATGATGATGAT 1506
DB 1450 ATCACTCTTAAGCCTTCCCTATTAAGCAAGCTTGTGATGATGCTGATGATGATGAT 1509
QY 1507 TTGGGTATCCGAAAGGAAACCAAGATGGGGCGAGTGAACCTTGTGATTAATCTGTTGAA 1566
DB 1510 TTGGGTATCCGAAAGGAAACCAAGATGGGGCGAGTGAACCTTGTGATTAATCTGTTGAA 1569
QY 1567 GTTGCCTATGTAAGCAATCCAGTCCCTATCACTAGCAAAAGCGACATATACCTTT 1626
DB 1570 GTTGCCTATGTAAGCAATCCAGTCCCTATCACTAGCAAAAGCGACATATACCTTT 1629
QY 1627 ACTGCAAGCGGCGGAGCAAGCTTGAATAATCTCCCTGATGATGATGATGATGATGATGAT 1686
DB 1630 ACTGCAAGCGGCGGAGCAAGCTTGAATAATCTCCCTGATGATGATGATGATGATGATGAT 1689
QY 1687 ACTACTGCTTCTGTAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746
DB 1690 ACTACTGCTTCTGTAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
QY 1747 ACAAGATATGTCGGAATGATCTTCAAGCAACATTTGCAATTAATGATGATGATGATGAT 1806
DB 1750 ACAAGATATGTCGGAATGATCTTCAAGCAACATTTGCAATTAATGATGATGATGATGAT 1809
QY 1807 AACGTAAGAAATGATTTATTAATTCGCGCAAGATGTAACATTAATGATGATGATGAT 1866
DB 1810 AACGTAAGAAATGATTTATTAATTCGCGCAAGATGTAACATTAATGATGATGATGAT 1869
QY 1867 GCATTAATGTCGCGGTGAGCAACAAATTTCTGTTGGCAATGTTGAATGATGAT 1920
DB 1870 GCATTAATGTCGCGGTGAGCAACAAATTTCTGTTGGCAATGTTGAATGATGAT 1923

```

```

RESULT 7
US-08-873-479-41
; Sequence 41, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Slome, Alan
; APPLICANT: Lynne, Christlanon
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3003 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-873-479-41

Query March 59.2%; Score 1136.8; DB 2; Length 3003;
Best Local Similarity 74.9%; Pred. No. 0;
Matches 1436; Conservative 0; Mismatches 477; Indels 3; Gaps 1;

QY 5 GAAAGAAAGAGTGTGTTTATCTGTTTATCAAGCTGACGATTCGTGCACTGTGCAT 64
DB 659 GATGGAAGAGGTTTTTTTATCCGTTTATCAAGTGTGCACTATGTCTTCTGTGCTT 918
QY 65 TAAACAAATCCCTCGGCTGTGATGCAAGACTTTTATCTGATTTTAAAGAAATTCAAA 124
DB 919 TAAACAGTCTTCTTATGATGAGGCGAAACATTTTGAATGGAATTTAAAGGGAATAGAA 978
QY 125 CAACACCAATGTCAGTGTCTTCTCCAAACAGCAACAAAGTGGCGGCTGCAATTTGCG 184
DB 979 CACTTAAGCTAGAGAGAGCTGCAACCAAGCAAGAAAGGGAAGGCAATCTTTTCTTG 1038
QY 185 TGGAGTGTGAAATGTGAAATCTTCTTAAAGATGCTTAAAGAACTTGAACAGTACCGG 244
DB 1039 TAAACCTGAAATGTGAAATCCCAAGAGTATTCAAAGAAAGTGAAGTGTGCTCAG 1098
QY 245 CAATTAATTAATCTCATATTTGTCATTAATTAATGAGCCCATTTTAAAGAAACAAACAGA 304
DB 1099 CGGATTAACAGCTATATATGCTTCAATTTGACGGAACCTATTTTAAAGAAACGCAATTC 1158
QY 305 AGCTAGAGCAACTGGAGCAAGATTTCTGACTACATCCCTGATTAAGCATATTTGTCG 364
DB 1159 AACTAGAGAGAGCGGAGCAAGAAATTTCTGATTTACATCAAGATTAAGCTTAATTTGCG 1218
QY 365 AGTATGAGGGAGATGTTCACTGCAAAAGTCCGCTCATTTGAACAGTGAATCAAGTGAAGC 424

```


Db 1219 AATATGATGGGGATGTAAAGCCGTAACATAACGAATTCGGCATTTTGAATCGGTTGAAC 1278
QY 425 CATCTTGGCCGAATACAAATATGATCCCAAGCTTTTTCACAAAAGCGCATGCAAGCTGG 484
Db 1279 CATATTACCTTTATATTAATTAAGACCCCGCAATTAATTTTTCAGAGAGGCTTCGAAATTAG 1338
QY 485 TGAAGCGTTGGCGCTTGATATACAGAGAGAAACAATAAGAGTGAATTAAGAGGATCG 544
Db 1339 TAGAAACAGTAGCTTTAGATTAAGAAAGCAAGAAAGTAAAGAGTACGTTTAAGAGATTGG 1398
QY 545 AGGAATTCGCTCGATAGCAAGCAATGACGTCATTAATATTAACGCAAAAGCCTGAAT 604
Db 1399 AACCAATGCCCAATACCGCAACAAATATGATATTAATGATACGTAACCCCAAGCCTGAAT 1458
QY 605 ATTAAGGTATGAAATGATGTGGCCAGAGATTTGTCAAGCGGATGTGGCAGAGACAGCT 664
Db 1459 ACGAATTTTGAATGACGTGGCCCGGTGCAATGTGAAAGCAGACGTCGCAAAAATTAATCT 1518
QY 665 ACGGTTTGTATGACAAAGGCAGATTTGTGCGAGTTGCGATCTGATGATTTGATATACAGAA 724
Db 1519 TTGGCTTAATATGACAAAGACAGATTTGTAGCAATTTGCTGATCTGGGCTTGATACAGAA 1578
QY 725 GAAACGACAGTTGATGATGATGAAGCCTTTCGCGGTAAATTAACGACATTAATGACATG 784
Db 1579 GAAATGACAGTTGATGATGATGAAGCCTTTCGCGGTAAATTAACGACATTAATGACATG 1638
QY 785 GTCCGACGAAATATGGAATGATGACAAAGGTATGTATCCCATGTGGCAGGTTCCGTAT 844
Db 1639 GCAAGAACGATTAACGCAATGATCCAAATGACATGAACCCCATGTTGCGATCTGTGT 1698
QY 845 TAGGAATGGCGCAACGAATTAAGAAATGGCACTCAAGCGAATCTGGTTTTCATTCGA 904
Db 1699 TAGGAAT---GCTAATAAATTAAGGATGGCAACCGCAACCAATCTAGTCTTTCATTCGA 1755
QY 905 TCATGATATGACAGTGTGGGCTTGAAGGCTTTCATCTGCAATCTGCAAACTTAATTCAGCC 964
Db 1756 TTATGATATGATGTGGAGGGCTGGAGGACATCTCTATATCAACAACTTAATTCAGTC 1815
QY 965 AAGCATTCAGTGCAGGTGCAGAAATTCATACAACTCTGGGGGGGCAAGCGGTGATGCG 1024
Db 1816 AAGCATATATGTCTGAGGCGAGAAATTCATACGAATTCATGCGGGGCTCCAGTAAACG 1875
QY 1025 CCTACACGACAGATTCAGAAATGTGATGACATATGAAGAAATATGATGACGATTC 1084
Db 1876 CCTATACGACAGCTTCGAAATGTGATGATTAATGAAGAAATATGATGACGATTC 1935
QY 1085 TTTTCGCGGCTGGAAATGAAGCCGACGAGCTGACATGATGACCTGATGACGCTTA 1144
Db 1936 TTTTTCGCGCCGGAATTAAGGACGATAGCGGTACATCACTGATGACAGAGAAACGCA 1995
QY 1145 AAAAGCGCAATACAGTGGCGCAACGAAACCTGCGTCAGAGCTTCGATTCATGCG 1204
Db 1996 AAAATGCGATTAAGTGGGGCAACGAAACCTAGTGCAGAGCTTCGATTCATGCG 2055
QY 1205 ATTAATATTAACACGCTTGAAGTTCCTTCCCTGGCCCGCAAAAGATGGCCGAATCA 1264
Db 2056 ATTAATATTAACACAGTGTGCTCAATTCCTTCACAGAGCTTCCTAAGATGACGATTA 2115
QY 1265 AGCTGATATGATGGCGGACGAGGACATACATTTATCAGACAGATCTTCTTTCGACCG 1324
Db 2116 AGCCGACGCTCATGAGACAGAGTACGTATATCTCTGCTAGATCATATTAAGCTCCAG 2175
QY 1335 ATTTCCTCTTCTGGCGGAATCATGACAGCAAAATATGCTTAATATGGGTGAACGTCATG 1384
Db 2176 ATTTCCTCATTTGGGGCAAAACATGATGATTAATATGCTTAATATGGGTGAACGTCATG 2235
QY 1385 CAACACCGATTTGTGGGGGAATGTTTGCACAGTCCGTACAGATTTTGTGAAGAAATAG 1444
Db 2236 CTACTTCGAATTTGAGAGGTAAATGTTGCACATTAAGGAGCAATTTTGTGAAGAAATAG 2295
QY 1445 GAATCACTCTTAAGCTTCCCTAATTAAGACAGCTTGAATGGAGGTGCTGATGATG 1504
Db 2296 GGGTAATCTCTAAGCTTCCCTTTTAAAGCTGCTTTAATTTGACAGGTGCTGCGAGATTTG 2355

QY 1505 GATTGGGTTATCCGAACGGAACCAAGATGGGGCCGAGTGAACCTGATTAATCGTTGA 1564
Db 2356 GACTTGGCTTTCCAAATGGTAAACAAAGATGGGGAAGATTAACGTTAATTAATCCCTTA 2415
QY 1565 ACGTGGCTATGTGAACGAATCAAGTCCCTATCACTAGCCAAAAAGGACATATACCT 1624
Db 2416 ATGTGCAATTTGTGAATGAAGAGAGCCCTTTTCAACAGTCAAAAGCAATATTCGT 2475
QY 1625 TTACTGCAACGGCGGCAAGCCATTGAAATCTCCCTGTATGTGCGATGCCCTGCA 1684
Db 2476 TTACGGCTCAAGCTGGTAAACCTTAAATATCACTTTTGTGTCAGATGACAGTA 2535
QY 1685 GCATCACTGCTCTGTAACCCCTGTCAATGATTTGATTTGATCAATACCAACCAAG 1744
Db 2536 GACAGCGGCAATCACTTAATTTAGTAATGATTTGATTAATCACTGACCAAAATG 2595
QY 1745 GAACAAAGATATGTGGGAATGACTTCTCAGACCAATTTGACATTAACGGAATGGCCGA 1804
Db 2596 GAATTAATATGCTCGAAATGACTTTACAGCACCGTATGATTAACATTTGGATGGCA 2655
QY 1805 ATTAAGTAATAATGATTTATTAATTCGCCCAAGTGAACATATACATTGAGTGC 1864
Db 2656 ACAAGTGAATAATGTTTATCAATGCTCTCAAGCGGAAGATACATGTCGAAGTGC 2715
QY 1865 AAGCATTAATATGCTCGGTTGACCAACAACTTCTGTTGGCAATTTGAACATA 1920
Db 2716 AAGCTTAATATGATCAAGTAAAGTCCGCAACCTTTTCTTACGATTTGATTA 2771

RESULT 8
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHRIFFLINGER, F.
APPLICANT: PALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0239
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: BP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 01:34:45, Search time 1378.42 Seconds
(without alignments)
11518.383 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atgagagaagaagaaggtgtt.....cgttgcaattgtgaactaa 1920

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA Main:*

1: /cgn2_6/prodata/1/pubpna/us07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/us08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/us09_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/us09B_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/us10A_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/us10B_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/us10C_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/us10D_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/us10E_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/us11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1920	100.0	1920	US-09-920-954-3	Sequence 3, Appl1
2	1920	100.0	1920	US-10-784-870-3	Sequence 3, Appl1
3	1477.2	76.9	1923	US-09-920-954-5	Sequence 5, Appl1
4	1477.2	76.9	1923	US-09-920-954-7	Sequence 7, Appl1
5	1477.2	76.9	1923	US-10-456-479-3	Sequence 3, Appl1
6	1477.2	76.9	1923	US-10-784-870-5	Sequence 5, Appl1
7	1477.2	76.9	1923	US-10-784-870-7	Sequence 7, Appl1
8	1477.2	76.9	1923	US-10-820-712A-2	Sequence 2, Appl1
9	1477.2	76.9	1923	US-10-820-712A-2	Sequence 2, Appl1
10	1295.6	67.5	1302	US-10-820-712A-11	Sequence 11, Appl1
11	1295.6	67.5	1302	US-10-820-712A-11	Sequence 11, Appl1
12	1002.6	52.2	1305	US-10-385-662-1	Sequence 12, Appl1
13	999.6	52.1	1302	US-10-820-712A-13	Sequence 13, Appl1
14	999.6	52.1	1302	US-10-820-712A-14	Sequence 14, Appl1
15	823.6	42.9	1302	US-10-820-712A-21	Sequence 21, Appl1
16	823.6	42.9	1302	US-10-820-712A-22	Sequence 22, Appl1
17	747.6	38.9	1299	US-10-820-712A-17	Sequence 17, Appl1
18	747.6	38.9	1299	US-10-820-712A-18	Sequence 18, Appl1
19	738	38.4	1299	US-10-820-712A-19	Sequence 19, Appl1
20	738	38.4	1299	US-10-820-712A-20	Sequence 20, Appl1
21	734.8	38.3	1299	US-10-820-712A-15	Sequence 15, Appl1
22	734.8	38.3	1299	US-10-820-712A-16	Sequence 16, Appl1
23	40	2.1	522	US-09-824-893A-56	Sequence 56, Appl1

24	40	2.1	522	3	US-09-824-893A-70	Sequence 70, Appl1
25	40	2.1	522	8	US-10-736-997-56	Sequence 56, Appl1
26	40	2.1	522	8	US-10-736-997-70	Sequence 70, Appl1
27	40	2.1	8911	3	US-09-070-927A-651	Sequence 651, Appl1
28	39.6	2.1	2382	5	US-10-084-817-312	Sequence 312, Appl1
29	39.6	2.1	2947	9	US-10-287-436A-213	Sequence 213, Appl1
30	39.6	2.1	4153	8	US-10-723-860-5175	Sequence 5175, Appl1
31	39	2.0	522	3	US-09-824-893A-50	Sequence 50, Appl1
32	39	2.0	522	8	US-10-736-997-50	Sequence 50, Appl1
33	38.6	2.0	1192	7	US-10-282-122A-35745	Sequence 35745, Appl1
34	38.4	2.0	522	3	US-09-824-893A-37	Sequence 37, Appl1
35	38.4	2.0	522	3	US-09-824-893A-47	Sequence 47, Appl1
36	38.4	2.0	522	3	US-09-824-893A-43	Sequence 43, Appl1
37	38.4	2.0	522	3	US-09-824-893A-60	Sequence 60, Appl1
38	38.4	2.0	522	3	US-09-824-893A-62	Sequence 62, Appl1
39	38.4	2.0	522	3	US-09-824-893A-66	Sequence 66, Appl1
40	38.4	2.0	522	3	US-09-824-893A-76	Sequence 76, Appl1
41	38.4	2.0	522	3	US-09-824-893A-79	Sequence 79, Appl1
42	38.4	2.0	522	3	US-09-824-893A-80	Sequence 80, Appl1
43	38.4	2.0	522	3	US-09-824-893A-83	Sequence 83, Appl1
44	38.4	2.0	522	3	US-09-824-893A-101	Sequence 101, Appl1
45	38.4	2.0	522	3	US-09-824-893A-112	Sequence 112, Appl1

ALIGNMENTS

```
RESULT 1
US-09-920-954-3
; Sequence 3, Application US/09920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKI, KATSUYOSHI
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; SOFTWARE: PatentIn version 3.0
; NUMBER OF SEQ ID NOS: 24
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
US-09-920-954-3
Query Match 100.0%; Score 1920; DB 3; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAAAAGAGAGAGTGTGTTTATCTGTTTATGCTGACGCACTTCTGCACTGT 60
DB 1 ATGAGAAAAGAGAGAGTGTGTTTATCTGTTTATGCTGACGCACTTCTGCACTGT 60
QY 61 GCATTAACAATCCCTCGCGCTGATGATGCAAGCACTTTGATGATGATTTAAAGAAAT 120
DB 61 GCATTAACAATCCCTCGCGCTGATGATGCAAGCACTTTGATGATGATTTAAAGAAAT 120
```

121 CAAACAAACCGATGTCAGTGGTTCTCCAAACAGCGACAAACAGGTGCGTGCATTT 180
121 CAAACAAACCGATGTCAGTGGTTCTCCAAACAGCGACAAACAGGTGCGTGCATTT 180
181 CTGGTGAAGCTGAAAATGTGAAACTTCTTAAAGATTGCTTAAAGAACTTGAACAGTA 240
181 CTGGTGAAGCTGAAAATGTGAAACTTCTTAAAGATTGCTTAAAGAACTTGAACAGTA 240
241 CCGGCAAAATATTAATCTCATAATGTCGAATTCANAGGCCCATTTTAAAGAAACAAA 300
241 CCGGCAAAATATTAATCTCATAATGTCGAATTCANAGGCCCATTTTAAAGAAACAAA 300
301 CAGAACTAGAGCAACTGAGCAAAAGATTCTGACTATCATCTGATTAATGATTAAT 360
301 CAGAACTAGAGCAACTGAGCAAAAGATTCTGACTATCATCTGATTAATGATTAAT 360
361 GTGAGTATGAGGGGAGATTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTG 420
361 GTGAGTATGAGGGGAGATTTCAGTCAAAAGTCCGCTCATTTGAACAGTGAATCAGTG 420
421 GAGCCATCTGCGCAATATCAAAATATGATCCCGAGCTTTTCAAAAAGCCGATCGACG 480
421 GAGCCATCTGCGCAATATCAAAATATGATCCCGAGCTTTTCAAAAAGCCGATCGACG 480
481 CTGGTGAAGCTGCGTGGCTGTGATACGAGCAAAACATATTAAGAGTGAATGAAGGC 540
481 CTGGTGAAGCTGCGTGGCTGTGATACGAGCAAAACATATTAAGAGTGAATGAAGGC 540
541 ATGAGAAATGCTCTCAGTACGTAGCAAGCAATGACGTCCATTAATTAACGCAAAAGCT 600
541 ATGAGAAATGCTCTCAGTACGTAGCAAGCAATGACGTCCATTAATTAACGCAAAAGCT 600
601 GAAATTAAGGTATGATGATGATGAGCAGAGGATTTGCAAAAGCGAGTGGCAAGAGC 660
601 GAAATTAAGGTATGATGATGATGAGCAGAGGATTTGCAAAAGCGAGTGGCAAGAGC 660
661 AGCTACGCTTGTATGAGCAAGCGCAAGATGTGCGAGTTCGCAATCTGATGATGATCA 720
661 AGCTACGCTTGTATGAGCAAGCGCAAGATGTGCGAGTTCGCAATCTGATGATGATCA 720
721 GGAAGAAACGACAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 780
721 GGAAGAAACGACAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 780
781 GGAAGAAACGACAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 780
781 GGAAGAAACGACAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATG 780
781 CTGGTGAAGCTGCGTGGCTGTGATACGAGCAAAACATATTAAGAGTGAATGAAGGC 840
781 CTGGTGAAGCTGCGTGGCTGTGATACGAGCAAAACATATTAAGAGTGAATGAAGGC 840
841 GTATTAGAAATGCGGCAACGAATTAAGGAATGCGACCTCAAGCGAATCTGTTTTTCA 900
841 GTATTAGAAATGCGGCAACGAATTAAGGAATGCGACCTCAAGCGAATCTGTTTTTCA 900
901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
961 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
961 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
961 AGCCAAAGCATTCAGTGAAGTTCGCAAAATCTCCTGAGGGGCGAGCGGTGAT 1020
961 AGCCAAAGCATTCAGTGAAGTTCGCAAAATCTCCTGAGGGGCGAGCGGTGAT 1020
1021 GGGGCTTACAGCAAGATTCGCAAAATGATGATGATGATGATGATGATGATGATGATG 1080
1021 GGGGCTTACAGCAAGATTCGCAAAATGATGATGATGATGATGATGATGATGATGATG 1080
1081 ATTCTTTTGGGCTGGGATGAAAGGCGGCAACGCGGTACATCAGTCACTGCTGATG 1140
1081 ATTCTTTTGGGCTGGGATGAAAGGCGGCAACGCGGTACATCAGTCACTGCTGATG 1140
1141 GCTAAATACGCAATTAAGTGGCGCAACGCAAACTGCGTCAAGCTGCTGCTGAT 1200
1141 GCTAAATACGCAATTAAGTGGCGCAACGCAAACTGCGTCAAGCTGCTGCTGAT 1200
1201 GCTAAATACGCAATTAAGTGGCGCAACGCAAACTGCGTCAAGCTGCTGCTGAT 1200
1201 GCTAAATACGCAATTAAGTGGCGCAACGCAAACTGCGTCAAGCTGCTGCTGAT 1200

1201 GCAATATATTAATTAACAGCTTGCACAGTCTCTCCGTGGCCCGCAAAAAGATGGCGGA 1260
1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
1261 ATCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
1321 CCGATTCCTCTCTGCGGCAATCATGACGCAAAATATGCTTATATGAGTGAACGCTCC 1380
1321 CCGATTCCTCTCTGCGGCAATCATGACGCAAAATATGCTTATATGAGTGAACGCTCC 1380
1381 ATGCAACACCGATTTGTTGCGGGAATGTTGCAAGCTCCGTGAGCAATTTTGAATAAT 1440
1381 ATGCAACACCGATTTGTTGCGGGAATGTTGCAAGCTCCGTGAGCAATTTTGAATAAT 1440
1441 AGAGAAATCACTCTTAAGCTTCCCTATGGAAGCACTTTGATGATGATGATGATGATG 1500
1441 AGAGAAATCACTCTTAAGCTTCCCTATGGAAGCACTTTGATGATGATGATGATGATG 1500
1501 GTTGAATGGGTTATCCGAACGGAACCAAGATGGGCGGAGTGAACCTGATTAATTCG 1560
1501 GTTGAATGGGTTATCCGAACGGAACCAAGATGGGCGGAGTGAACCTGATTAATTCG 1560
1561 TTGAACCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1561 TTGAACCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1621 ACCTTATCTGCAACGCGGCGGCAAGCAATGGAATCTTCTGATGATGATGATGATGATG 1680
1621 ACCTTATCTGCAACGCGGCGGCAAGCAATGGAATCTTCTGATGATGATGATGATGATG 1680
1681 GCAAGCACTACTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1681 GCAAGCACTACTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1741 AACGAAACAGATATGTCGGAATGATGATGATGATGATGATGATGATGATGATGATG 1800
1741 AACGAAACAGATATGTCGGAATGATGATGATGATGATGATGATGATGATGATGATG 1800
1801 CGCAATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
1801 CGCAATTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
1861 GTGCAACATATATGTCGCGGTTGAGCAAAACCTTCTGTTGCAATTTGATGATGATG 1920
1861 GTGCAACATATATGTCGCGGTTGAGCAAAACCTTCTGTTGCAATTTGATGATGATG 1920

RESULT 2
US-10-784-870-3
Sequence 3, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0BCT
CURRENT APPLICATION NUMBER: US/10/784, 870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 3
/ LENGTH: 1920
/ TYPE: DNA
/ ORGANISM: Bacillus sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1920)
US-10-784-870-3

Query Match 100.0%; Score 1920; DB 7; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATACGCTGACCGCATTCGTGACCTGT 60
DB 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATACGCTGACCGCATTCGTGACCTGT 60
QY 61 GCATTAAACATATCCCTGCGCTGGTGAAGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
DB 61 GCATTAAACATATCCCTGCGCTGGTGAAGCAAGGACTTTGATCTGATTTTAAAGAAAT 120
QY 121 CAAACAAACACCGATGTCAGTGTGTTCTCAACACGCGCAAAACAGTGGCGCTGCATT 180
DB 121 CAAACAAACACCGATGTCAGTGTGTTCTCAACACGCGCAAAACAGTGGCGCTGCATT 180
QY 181 CTGGTGAAGTCTGAAAATGTGAACCTTTTAAAGATGCTTAAAGAACTTTGAAACAGTA 240
DB 181 CTGGTGAAGTCTGAAAATGTGAACCTTTTAAAGATGCTTAAAGAACTTTGAAACAGTA 240
QY 241 CCGGCAATATATTAATCTCATATGTCATATTCATATGCGCCCATTTTAAAGAAACAAA 300
DB 241 CCGGCAATATATTAATCTCATATGTCATATTCATATGCGCCCATTTTAAAGAAACAAA 300
QY 301 CAGAGCTAGAGCAACCTGAGCAAAAGATTCCTGACTACATCCCTGATTAATGATATAT 360
DB 301 CAGAGCTAGAGCAACCTGAGCAAAAGATTCCTGACTACATCCCTGATTAATGATATAT 360
QY 361 GTTCAGTATGAGGGGGATGTTCAGTCAAAAGTCCTGCTCAATGGAACAATGGAATGAG 420
DB 361 GTTCAGTATGAGGGGGATGTTCAGTCAAAAGTCCTGCTCAATGGAACAATGGAATGAG 420
QY 421 GAGCCATATCTGCGCAATATACAAATATGATCCCACTTTTCAAAAAGCGCATGAG 480
DB 421 GAGCCATATCTGCGCAATATACAAATATGATCCCACTTTTCAAAAAGCGCATGAG 480
QY 481 CTGGTGAAGCGTGGCGCTTGAACGAGCAAGAACATTAAGAGTGAATTAAGAGC 540
DB 481 CTGGTGAAGCGTGGCGCTTGAACGAGCAAGAACATTAAGAGTGAATTAAGAGC 540
QY 541 ATCGAGGAATCGCTCAGTACGTAGCAAGCAATGACCTCATTAATTAACGCAAGCT 600
DB 541 ATCGAGGAATCGCTCAGTACGTAGCAAGCAATGACCTCATTAATTAACGCAAGCT 600
QY 601 GAAATTAAGGATATGATGATGAGCAGAGGATTTGTAAGCGGATGTCACAGAC 660
DB 601 GAAATTAAGGATATGATGATGAGCAGAGGATTTGTAAGCGGATGTCACAGAC 660
QY 661 AGCTACGTTTGTATGAGCAAGCGCATTTGTCAGCTTCCGATCTGATTTGATACA 720
DB 661 AGCTACGTTTGTATGAGCAAGCGCATTTGTCAGCTTCCGATCTGATTTGATACA 720
QY 721 GGAAGAAACGACGTTGATGATGAGAGCTTCCGCGTAAATTAACAACATATATGCA 780
DB 721 GGAAGAAACGACGTTGATGATGAGAGCTTCCGCGTAAATTAACAACATATATGCA 780
QY 781 CTGGGTGGAAGAAATATGAGATGATGAGAGCGTCAATGATCCCATGAGGATG 840
DB 781 CTGGGTGGAAGAAATATGAGATGATGAGAGCGTCAATGATCCCATGAGGATG 840
QY 841 GTATTAGAAATGCGCGCAAGATTAAGAAATGCGCACTCAAGCGAATCTGTTTTTCA 900
DB 841 GTATTAGAAATGCGCGCAAGATTAAGAAATGCGCACTCAAGCGAATCTGTTTTTCA 900
```

```
QY 901 TCCATCATGATATAGCAGTGTGGGCTTGGAGGCTTGCTTCCATCTGAAACCTTATTC 960
DB 901 TCCATCATGATATAGCAGTGTGGGCTTGGAGGCTTGCTTCCATCTGAAACCTTATTC 960
QY 961 AGCAGACATTCAGTGCAGGTGCAGAAATTCATACAACTCTGGGGGCGACGGTGAAT 1020
DB 961 AGCAGACATTCAGTGCAGGTGCAGAAATTCATACAACTCTGGGGGCGACGGTGAAT 1020
QY 1021 GGGGCTTACACACAGATTCAGAAATGTGATGACTATGTAAGAAAAATGATATAGC 1080
DB 1021 GGGGCTTACACACAGATTCAGAAATGTGATGACTATGTAAGAAAAATGATATAGC 1080
QY 1081 ATTCTTTTGGCGCTGGGAAATGAAAGCGCGGATACCATGATGCACTGTGATG 1140
DB 1081 ATTCTTTTGGCGCTGGGAAATGAAAGCGCGGATACCATGATGCACTGTGATG 1140
QY 1141 GCTTAAAAAGCCATTAACAGTCCGCGCAACGAAACCTGCTCAAGCTTCGGTCTTAT 1200
DB 1141 GCTTAAAAAGCCATTAACAGTCCGCGCAACGAAACCTGCTCAAGCTTCGGTCTTAT 1200
QY 1201 GCAGATTAATATTAACCAAGTTCACAGTTCCTTCCGTCGCGCAAAAGATGGCGCA 1260
DB 1201 GCAGATTAATATTAACCAAGTTCACAGTTCCTTCCGTCGCGCAAAAGATGGCGCA 1260
QY 1261 ATCAAGCTGATGTCAATGGCGCGCAGGAGATATATTTATCAGCAAGATCTTCTTGA 1320
DB 1261 ATCAAGCTGATGTCAATGGCGCGCAGGAGATATATTTATCAGCAAGATCTTCTTGA 1320
QY 1321 CCCGATTCCTCTCTGCGGCAATCAATGACAGAAATATGCTATATGAGTGAAGCTGC 1380
DB 1321 CCCGATTCCTCTCTGCGGCAATCAATGACAGAAATATGCTATATGAGTGAAGCTGC 1380
QY 1381 ATGGCAACACCGATTTGTGCGGGGAATGTGCAACAGCTCCGTGAGCATTTTGTGAAAAT 1440
DB 1381 ATGGCAACACCGATTTGTGCGGGGAATGTGCAACAGCTCCGTGAGCATTTTGTGAAAAT 1440
QY 1441 AAGAGATATATCTCTTAAAGCTTCTTATTTGAAAGCAGCTTTGATGCAAGTGTGAT 1500
DB 1441 AAGAGATATATCTCTTAAAGCTTCTTATTTGAAAGCAGCTTTGATGCAAGTGTGAT 1500
QY 1501 GTTGAATGGGTTATCCGAACGGAATGAGGAGCGGATGACCTGTGATTAATG 1560
DB 1501 GTTGAATGGGTTATCCGAACGGAATGAGGAGCGGATGACCTGTGATTAATG 1560
QY 1561 TTGAACGTTGCTTATGTGAACGAATCAAGTGCCTATGAATGCAAAAAGCAATAT 1620
DB 1561 TTGAACGTTGCTTATGTGAACGAATCAAGTGCCTATGAATGCAAAAAGCAATAT 1620
QY 1621 ACCTTTTATGCAACGCGCGGCAAGCATTTGAAAATCTCCCTGTATGTGCGATGCCCT 1680
DB 1621 ACCTTTTATGCAACGCGCGGCAAGCATTTGAAAATCTCCCTGTATGTGCGATGCCCT 1680
QY 1681 GGAAGCACTACGCTTCTGTAACCTGTGCAATGATTTGATTTGATGATTAACAGACA 1740
DB 1681 GGAAGCACTACGCTTCTGTAACCTGTGCAATGATTTGATTTGATGATTAACAGACA 1740
QY 1741 AACGAAACAAGATATGTGCGGATGACTTCTCAGACCAATTTGACAATTAATCTGGATG 1800
DB 1741 AACGAAACAAGATATGTGCGGATGACTTCTCAGACCAATTTGACAATTAATCTGGATG 1800
QY 1801 CGCAATTAACGTAAGAAATATGATTTATTAATGCGCCCAAGTGAACATATACATTTGAG 1860
DB 1801 CGCAATTAACGTAAGAAATATGATTTATTAATGCGCCCAAGTGAACATATACATTTGAG 1860
QY 1861 GTGCAAGATATTAATGTGCGGTTGGAACCAAAAATCTCTGTGTCGAATGTGAATCTTA 1920
DB 1861 GTGCAAGATATTAATGTGCGGTTGGAACCAAAAATCTCTGTGTCGAATGTGAATCTTA 1920
```

RESULT 3
US-09-920-954-5
; Sequence 5. Application US/09920954
; Publication No. US20020064854A1

GENERAL INFORMATION:
APPLICANT: TAKAIMA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920,954
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-5

Query Match 76.9%; Score 1477.2; DB 3; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1441; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 7 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGCGATTCTGTGACATGTTGCATTA 66
DB 10 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGCGATTCTGTGACATGTTGCATTA 69
QY 67 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGCGATTCTGTGACATGTTGCATTA 126
DB 70 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGCGATTCTGTGACATGTTGCATTA 129
QY 127 AACAACGATGTGAGTGTGTTTCTCAAAACGACGACAAACAGTGTGCGATTTCTGTG 186
DB 130 AACAACGATGTGAGTGTGTTTCTCAAAACGACGACAAACAGTGTGCGATTTCTGTG 189
QY 187 GAGTGTGAAGATGTGAACCTTCTTAAAGGATGCTTAAAGAAACCTTGAACAGTACGCGCA 246
DB 190 GAGTGTGAAGATGTGAACCTTCTTAAAGGATGCTTAAAGAAACCTTGAACAGTACGCGCA 249
QY 247 AATAATAAATCCCATATGTCATTCATTCATTCGCTTTCATTCATTCGCTTTCATTCATTC 306
DB 250 AATAATAAATCCCATATGTCATTCATTCATTCGCTTTCATTCATTCGCTTTCATTCATTC 309
QY 307 CTAGAGCAACTGAGCAAAAGATCTCGACATACCTCCGATTATGATATGTCGAG 366
DB 310 CTAGAGCAAAAGATCTCGACATACCTCCGATTATGATATGTCGAG 369
QY 367 TATGAGGGGAGTGTGAGTCAAAAGTCCGCTTCATTTGAACAGTGAAGCA 426
DB 370 TATGAGGGGAGTGTGAGTCAAAAGTCCGCTTCATTTGAACAGTGAAGCA 429
QY 427 TACTTCCGGAATTAATAATGATTCCTCCAGCTTTTCAAAAAGGCGCATCCAGTGTG 486
DB 430 TACTTCCGGAATTAATAATGATTCCTCCAGCTTTTCAAAAAGGCGCATCCAGTGTG 489
QY 487 AAGAGTGTGCGCTGATACGAGCAAGCAATTAAGAAGTCAATTAAGAAGCTGAG 546
DB 490 AAGAGTGTGCGCTGATACGAGCAAGCAATTAAGAAGTCAATTAAGAAGCTGAG 549
QY 547 GAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
DB 550 GAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609

QY 607 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
DB 610 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
QY 667 GATTGTGATGACAAAGCCAGATTTGTCAGTTCCTGATTCGATTCGATTCGATTCGATTCGAT 726
DB 670 GATTGTGATGACAAAGCCAGATTTGTCAGTTCCTGATTCGATTCGATTCGATTCGATTCGAT 729
QY 727 AAGCAAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
DB 730 AAGCAAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
QY 787 CGAGCAATTAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
DB 790 CGAGCAATTAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
QY 847 GGAATGCGCGAAGCAATTAAGAATGCGCACTCAAGCAATCTGTTTCAATCCATC 906
DB 850 GGAATGCGCGAAGCAATTAAGAATGCGCACTCAAGCAATCTGTTTCAATCCATC 909
QY 907 ATGATAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 966
DB 910 ATGATAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 969
QY 967 GCATTCAGTGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1026
DB 970 GCATTCAGTGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1029
QY 1027 TACACAGCAGATTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
DB 1030 TACACAGCAGATTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
QY 1087 TTCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1146
DB 1090 TTCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1149
QY 1147 AAGCAGTAAAGTGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1206
DB 1150 AAGCAGTAAAGTGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1209
QY 1207 AATAATAAATCCCATATGTCATTCATTCATTCGCTTTCATTCATTCGCTTTCATTCATTC 1266
DB 1210 AATAATAAATCCCATATGTCATTCATTCATTCGCTTTCATTCATTCGCTTTCATTCATTC 1269
QY 1267 CCGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
DB 1270 CCGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
QY 1327 TCCCTCTTCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1386
DB 1330 TCCCTCTTCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1389
QY 1387 AACCGAATGTTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1446
DB 1390 AACCGAATGTTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1449
QY 1447 ATCACTCTTAAGCTTCCCTTAATGAAGCAGCTTTGATGAGTGTGATGAGTGTGATGAGTGTGAT 1506
DB 1450 ATCACTCTTAAGCTTCCCTTAATGAAGCAGCTTTGATGAGTGTGATGAGTGTGATGAGTGTGAT 1509
QY 1507 TTTGGTTATCCGGAACGGAACCAAGATGCGGCGGAGTGAACCTCGGATTAATGCTTGAAC 1566
DB 1510 TTTGGTTATCCGGAACGGAACCAAGATGCGGCGGAGTGAACCTCGGATTAATGCTTGAAC 1569
QY 1567 GTTCCCTATGTAAGCAATCAAGTCCCTATCACTAGCCAAAAGGACATATACCTTT 1626
DB 1570 GTTCCCTATGTAAGCAATCAAGTCCCTATCACTAGCCAAAAGGACATATACCTTT 1629
QY 1627 ACTGCAACGCGCGCAAGCATTTGAATAATCTCCCTGTATGCTGTGATGCTGTGCAAGC 1686
DB 1630 ACTGCAACGCGCGCAAGCATTTGAATAATCTCCCTGTATGCTGTGATGCTGTGCAAGC 1689

Qy 1387 ACACGATTGTTGGGGGATGTTGCACAGCTCCGAGCAATTTTGTGAAAAATAGAGA 1446
Db 1390 ACACGATCGTTGCGGAAACGTGGCAAGCTTGGAGCATTTTGTGAAAAACAGAGC 1449
Qy 1447 ATCACTCTTAAGCCTTCCCTATTTGAAGACGCTTGAATGCGAGTGCTGCTGATGTTGA 1506
Db 1450 ATCAACACCAAGCCTTCTATTTAAAGCGGCACTGATTTGCCGGTGCAGCTGACATGCGC 1509
Qy 1507 TTGGTTATCCGAACGGAACCAAGATGGGGCCGAGTGAACCTTGATTAATCGTTGAAC 1566
Db 1510 CTGGCTACCCGAAACGTTAACCAAGATGGGGCAAGTGAACATTTGATTAATCCCTGAAC 1569
Qy 1567 GTTGCTATGTAAGCAATCCAGTGCCTATCAATAGCAAAAGCAATATACCTTT 1626
Db 1570 GTTGCTATGTAAGCAATCCAGTGCCTATCAATAGCAAAAGCAATATACCTTT 1629
Qy 1627 ACTGCAACGCGCGGCAAGCCATTTGAAAAATCTCCCTGATGATGCGATGCCCTGCAAGC 1686
Db 1630 ACTGCTACGCGCGGCAAGCCCTTTGAAAAATCTCCCTGATGATGCGATGCCCTGCAAGC 1689
Qy 1687 ACTACTGCTTCTGTATACCTTGATGATGATTTGATTTGATTTGATTTGATTTGATTTG 1746
Db 1690 ACAACTGCTTCTGTATACCTTGATGATGATTTGATTTGATTTGATTTGATTTGATTTG 1749
Qy 1747 ACAAGATATGTCGGGATGATCTTCTCAGACCACTTTGCAATTAATCTGGGATGGCGCAAT 1806
Db 1750 ACAAGATATGTAAGCAATGATCTTCTCAGACCACTTTGCAATTAATCTGGGATGGCGCAAT 1809
Qy 1807 AACGTAGAAAAATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1866
Db 1810 AACGTAGAAAAATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1869
Qy 1867 GCATTTATGTCGGGTTGGACCAAAAATCTTCTGTTGGCAATTTGATTTA 1920
Db 1870 GCATTTATGTCGGGTTGGACCAAAAATCTTCTGTTGGCAATTTGATTTA 1923

RESULT 5

US-10-456-479-3
; Sequence 3, Application US/10456479
; Publication No. US20040072321A1
; GENERAL INFORMATION:
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: TAKIMURA, YASUHI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: NOMURA, MASAFUMI
; APPLICANT: KOBAYASHI, TOHRU
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 238700US0
; CURRENT APPLICATION NUMBER: US/10/456,479
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: JP 2002-186387
; PRIOR FILING DATE: 2002-06-26
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
; OTHER INFORMATION:
US-10-456-479-3

Query Match 76.9%; Score 1477.2; DB 7; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 7 AAGAGAGGTGTTTTTATCTGTTTTATCAGTCGAGGATTTCTGTCAGTGTGATTA 66
Db 10 AAGAAAAAGGTGTTTTTATCTGTTTTATCAGTCGAGGATTTCTGTCAGTGTGATTA 69
Qy 67 AACATCCCTGGGCTGTGATGCAAGGACTTTGATCTGATTTTAAAGAAATTCAAACA 126
Db 70 AGTATCCATCTGCAAGGTGTGCAAGGAAATTTGATCTGATTTTCAAGAAATTCAGACA 129
Qy 127 ACAACCAATGTCAGTGTGTTTTCCAAACAGGAAACAAACAGTGGGCTGCAATTTCTGTTG 186
Db 130 ACAACCAATGTCAGTGTGTTTTCCAAACAGGAAACAAACAGTGGGCTGCAATTTCTGTTG 189
Qy 187 GAGTCTGAATATGTAACCTTTAAAGATTTGCTTAAAGAACTTGAACAGTACCGGCA 246
Db 190 GAGTCTGAATATGTAACCTTTAAAGATTTGCTTAAAGAACTTGAACAGTACCGGCA 249
Qy 247 AATATTAACCTCAATTTGTCATTAATGAGCCCATTTTGAAGAAACAAACAGAG 306
Db 250 AATATTAACCTCAATTTGTCATTAATGAGCCCATTTTGAAGAAACAAACAGAG 309
Qy 307 CTAGAGCACTGAGCAAGAAATTTCTGACTCAATCCCTGATTTATGATATATTTGAG 366
Db 310 CTAGAGCACTGAGCAAGAAATTTCTGACTCAATCCCTGATTTATGATATATTTGAG 369
Qy 367 TATGAGGGGATGTTCAAGTCAAAAGTCCGCTCAATTGAACAGTGAATCAGTGAAGCA 426
Db 370 TATGAGGGGATGTTCAAGTCAAAAGTCCGCTCAATTGAACAGTGAATCAGTGAAGCA 429
Qy 427 TACTTCCGAAATACAAATATGATCCCAAGTTTTCACAAAGGCGCATGACCTGAGT 486
Db 430 TACTTCCGAAATACAAATATGATCCCAAGTTTTCACAAAGGCGCATGACCTGAGT 489
Qy 487 AAGGTTGGGCTGATTAAGCAAGCAATTAAGAGCAATTAAGGGATGAG 546
Db 490 AAGGTTGGGCTGATTAAGCAAGCAATTAAGAGCAATTAAGGGATGAG 549
Qy 547 GAAATCCCTCAGTACGATGCAAGCAATGACCTTCATTTATTTACGCAAAAGCTGAAAT 606
Db 550 GAAATCCCTCAGTACGATGCAAGCAATGACCTTCATTTATTTACGCAAAAGCTGAAAT 609
Qy 607 AAGGTATGAATGATGTCGCAAGGATTTGTCAAAGCGATGTCGACAGAGCACTAC 666
Db 610 AAGGTATGAATGATGTCGCAAGGATTTGTCAAAGCGATGTCGACAGAGCACTAC 669
Qy 667 GCTTGTATGCAAGGCAAGGATTTGTCCAGTTCGATTTGATTTGATTTGATTTGATTTG 726
Db 670 GCTTGTATGCAAGGCAAGGATTTGTCCAGTTCGATTTGATTTGATTTGATTTGATTTG 729
Qy 727 AACGACAGTTCGATGATGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGT 786
Db 730 AACGACAGTTCGATGATGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGT 789
Qy 787 CGAGCAATTAATGCAATGATGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGT 846
Db 790 CGAGCAATTAATGCAATGATGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGT 849
Qy 847 GGAATATGCGCAAGAAATTAAGAAATGCAAGGCTTCGCGGTAAATTAACGACATATATG 906
Db 850 GGAATATGCGCAAGAAATTAAGAAATGCAAGGCTTCGCGGTAAATTAACGACATATATG 909
Qy 907 ATGATAGCAGTGTGAGGCTTTGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGT 966
Db 910 ATGATAGCAGTGTGAGGCTTTGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGT 969
Qy 967 GCATTCAGTGTGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGTAAATGAGGCT 1026
Db 970 GCATTCAGTGTGAGGCTTCGCGGTAAATTAACGACATATATGCACTGGGTAAATGAGGCT 1029
Qy 1027 TACACGACATTTCCAAATATGTCAGTATGATTTGATTTGATTTGATTTGATTTGATTTG 1086
Db 1030 TACACGACATTTCCAAATATGTCAGTATGATTTGATTTGATTTGATTTGATTTGATTTG 1089
Qy 1087 TTCGCGCTGGGAATGAAGGCGCAAGGCGGTACCATTCAGTGCACCTGTATCGGCTTAA 1146

QY	787	CGAGCAATAAATGCCAATGATATACGAACGGTCAATGTAACCATGTGGCAGGTTCCGATTTA	846
Db	790	CGAGCAATAAATGCCAATGATATGCAATGTGCATGTATACGATGTGGTGGCTCCGATTTA	849
QY	847	GGAAATATGGCGCAACGAAATAAAGGAATGGGACCTCAAGCGAATCGGTGTTTCAATCCATC	906
Db	850	GGAAACGGCTCCACTATATAAGGAATGGCGCTCAAGCGAATCTAGCTTCCATCTATC	909
QY	907	ATGATATAGCAGTGTGTGGGCTTGGAGGCTTGCTCCATCTGCAAACTTATTCAGCCAA	966
Db	910	ATGATATAGCGGTGTGGGGAATTGGAGGACTATCTTCGATATGGAACCTTATTCAGCCAA	969
QY	967	GCAATTCAGTGCAGAGTCCAGAAATTCATACAACTCTGSGGGGCAAGCGGTGAATGGGCTC	1020
Db	970	GCAATCAGTGCAGTGTGTCCAGAAATTCATACAACTCTGSGGGGACAGCAGTGAATGGGCTC	1020
QY	1027	TACACGACGAGATTCACGAATATGTGTGATGTACTATGTAAAGAAAATGATATAGAGATTTCT	108
Db	1030	TACACGACGAGATTCACGAATATGTGTGATGTACTATGTGTGCAAAAATGATATAGAGATTTCT	108
QY	1087	TTTCGGGCTGGGAATGAAGCGCGAAGCGGCGGTACATCAGTGCACCTGGTACGGCTAAA	114
Db	1090	TTTCGGGCTGGGAATGAAGCGCGAAGCGGCGGTACATCAGTGCACCAAGCAAGCTAAA	114
QY	1147	AAAGCCATACAGTGTGGCGCAACCGAAAACCTGCGTCAAGCTTCGATTCATATGCAAT	120
Db	1150	AATGCAATTAACAGTGTGGCACTACGAAAACCTTCGCGCCAAAGCTTGTGGTCTTATGTGGGAC	120
QY	1207	AATATTAACAGGTGTGCAAGTTCCTTCGCGGGCGGCAAAAGATGGGCGAATCAAG	126
Db	1210	AATATTAACAGTGTGCAAGTTCCTTCGCGGGCGGCAAAAGATGGGCGAATCAAG	126
QY	1267	CTGTATGTATGAGCGCGCAAGGACATACATTTTATCAGCAAGATCTTCTCTGCAACCGAT	132
Db	1270	CGGATGTATGTGGCAACGGGAAGCTTCAATATACAGAAAGCTTCTCTGCAACCGGAT	132
QY	1327	TTCTCTCTTGTGGCGGAATCATGACAGCAAAATATGCTATATGTGGGTGAACGTCCATGGCA	138
Db	1330	TTCTCTCTTGTGGCGGAACCATGACAGTAAATATGCAATAGGTGTGAACGTCCATGGCT	138
QY	1387	ACACGGAATGTGGCGGGGAATGTGTGACAGCTCCGAGACATTTGTGAAAAATAGAGA	144
Db	1390	ACACGGAATGTGTGGGAACGTGGGACAGCTTCGAGGATTTGTGAAAAACAAGG	144
QY	1447	ATCACTCTTAAGCCTTCCCTATTTGAAAGAGCTTTGATTTGACAGTCTGCTGATGTTGGA	150
Db	1450	ATCAACACMAAACCTTCTCTATTTAAAGGCGGACTATGTTCGCGGTGACGTGACATCGGC	150
QY	1507	TTGGGTATTCGAGCGGAACCAAGATGGGCGGAGTACCCTGTGATTAATGTTGAAC	156
Db	1510	CTTGGCTATCCGAACGGTAAACCAAGATGGGCGGAGTACATTTGATTAATCCCTGAAC	156
QY	1567	GTGTGCTATGTGAAGAAATCCAGTGTGCCATATGAACCTAGGCCAAAACCGACATATACCTT	162
Db	1570	GTGTGCTATGTGAAGAGTCCAGTGTCTATATCAACAGGCCAAAACCGAGTATCTGTT	162
QY	1627	ACTGCAAGGCGGCGAGCAGCATTTGAAAAATCTCCCTGTATGTGTGCGATGCCCTGCAAG	168
Db	1630	ACTGTATCTGCGGCGAGCAGCTTTGAAAAATCTCCCTGTATGTGTGATGTCCTCGGAGC	168
QY	1667	ACTACTGCTTCTGTAAACCTGTGATCAATGATTTGGATTTGTGCTATTAACAGCACMAACGA	174
Db	1690	ACAATGTCTTCGTAACGCTTGTCAATGATCTGGAACCTTGTCAATACCGCTCCAAATGAG	174
QY	1747	ACAAAGATATGTGGGAATGAACCTTTCAGACCACTTTGACAAATATACCTGGGATGGCCGCAAT	180
Db	1750	ACACAGTATGTAGGAATGAACCTTTACTTGTGCAATACAAAGATATACCTGGGATGGCCGCAAT	180
QY	1807	AAAGTAAAGAAATGTATTTATTAATTTGCGCCCAAGTGAACATATACATTCGATGAGGTGCA	186
Db	1810	AAAGTAAAGAAATGTATTTATTAATGACCAACAAAGGCGGAGCTGTATCAATTTGAGGTACAG	186
QY	1867	GCATTAATATGTGCGGATTTGACCAACAACTTCTCGTTGGCAATGTGACTAA	1920

DB	1870	GCCTATAACGTAACCGGTTGACACCAAGACCTTCTGTTGGCAATTGTGAATTAA	1923
		<pre> RESULT 7 US-10-784-870-7 ; Sequence 7, Application US/10784870 ; Publication No. US20040142837A1 ; GENERAL INFORMATION: ; APPLICANT: TAKAIMA, MIKIO ; APPLICANT: OKUDA, MITSUYOSHI ; APPLICANT: SAEKI, KATSUHIISA ; APPLICANT: KUBOTA, HIROMI ; APPLICANT: HITOMI, JUN ; APPLICANT: KAGEYAMA, YASUSHI ; APPLICANT: SHIKATA, SHITSUM ; APPLICANT: NOMURA, MASAFUMI ; TITLE OF INVENTION: ALKALINE PROTEASE ; FILE REFERENCE: 0327-0832-0PCT ; CURRENT APPLICATION NUMBER: US/10/784,870 ; CURRENT FILING DATE: 2004-02-24 ; PRIOR APPLICATION NUMBER: US/09/509,814A ; PRIOR FILING DATE: 2000-04-06 ; PRIOR APPLICATION NUMBER: PCT/J998/04528 ; PRIOR FILING DATE: 1998-10-07 ; PRIOR APPLICATION NUMBER: JP 9-274570 ; PRIOR FILING DATE: 1997-06-08 ; NUMBER OF SEQ ID NOS: 24 ; SOFTWARE: PatentIn version 3.0 ; SEQ ID NO 7 ; LENGTH: 1923 ; TYPE: DNA ; ORGANISM: Bacillus sp. ; FEATURE: ; NAME/KEY: CDS ; LOCATION: (1)..(1923) US-10-784-870-7 </pre>	
QY	Query Match	76.9%; Score 1477.2; DB 7; Length 1923;	
Db	Best Local Similarity	85.7%; Pred. No. 0;	
	Matches 1641; Conservative	0; Mismatches 273; Indels 0; Gaps	0
QY	7	AAGAGAAAGGTGTTTATCTGTTTATCAGCTGACGCGATTCTGTCAGTGTTCATTA	66
Db	10	AAGAAAAAGGTGTTTATCTGTTTATCAGCTGACGCGATTGTGTCAGTGTTCGTTA	69
QY	67	AACAACTCCCTGGCGTGGTGAATGCAAGACCTTTGATCTGCAATTTAAAGAAATTAACA	126
Db	70	AGTATTCATCTGACGGTGGTGCAGAAATTTGATCTGCAATTTCAAGAAATTCAGACA	129
QY	127	ACAACCGATGTCAGTGGTTCCTCCAAACAGGCAACAAACAGGTCGGCTGCATTTCTGGTG	186
Db	130	ACAACGTATGCTTAAAGGTTTCTCCAAAGGAGGCGAGCTGGTGTCTGCTTTTCTGGTG	189
QY	187	GAGTCTGAAAATGTGAAACCTTCTTAAAGATTGCTTAAAGAACTTGAACAGTACCGGCA	246
Db	190	GAATCTGAAAATGTGAAACCTCCAAAGGTTTGCAGAAAGAGCTTGAACAGTCCGGCA	249
QY	247	AATATAAATCTCCATATTGTGCAATTCATAGGCCCATTTTAAAGAAACAAACAGAG	306
Db	250	AATATAAATCTCCATATTATCAATTCATAGGACCAATTTTAAAGAAACAAACAGAG	309
QY	307	CTAAGAGCAATGTGAGCAAGAAAGTTCGACCTACATCCCTGATTTATGATATTATGTGGAG	366
Db	310	CTGAAAAAACAAGGGCGAAAGATTTCTGACCTACATCCTGATTTATGCTTACATGTGAG	369
QY	367	TATGAGGGGAGATGTTCACTCAAAAGTCGCTCCATTGAAACAGTGAATCAGTGAAGCCA	426
Db	370	TATAGGGGCGATGTTAGTACGAAACAAGACACATTAGACAGGTGAGATCCGTGAGGCT	429
QY	427	TACTTGCCGAATAACAAATAGATCCCAAGCTTTTCAAAAGGCGCATGACGCTGGTG	486
Db	430	TATTTGCCGATATACAGAAATGATCCCAAGCTTTTCAAAAGGGGCATTCAGAGCTTGTA	489

130 ACAAGTATGCTAAAGTTTCTCCAGACGAGCGGACGACTGCTGCTCTTTCTGGTG 189
187 GAGTGTAAATGTGAACTTCTTAAGAGTTGCTAAAGAACTTGAACAGTACCGCA 246
190 GAATCTGAAATGTGAACTCCCAAGAGTTGCAAGAAAGCTTGAACAGTCCCGCA 249
247 AATATTAATCCCAATGTGCTCAATTCATAGGCCCCATTTTGAAGAAACAAACAGAG 306
250 AATATTAATCTCAATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 309
307 CTAGACACACTGAGCAAGATTTCTGACTATCATCTGATTAATATTTGTCAG 366
310 CTGAAAAACAGGGGCAAGATTTCTGACTATCATCTGATTAATGCTTACATTTGTCAG 369
367 TATGAGGGGAGTTTCACTCAAAAGTCCGCTCATTTGAAACAGTGAATCAGTGAAGCCA 426
370 TATGAGGGGAGTTTCACTCAAAAGTCCGCTCATTTGAAACAGTGAATCAGTGAAGCCA 429
427 TACTTGGCGAAATCAAAATAGATGCTCCGACTTTTCAAAAAGGGGCAATCCAGCTGTG 486
430 TATTTGCGATTAACAGATATGATCCGAGCTTTTCAAAAAGGGGCAATCCAGCTGTG 489
487 AAAGCTTGGCGCTTGTATACGAGCAAGCAATTAAGAGTGCATTAAGAGGCAATCGAG 546
490 AAAGCAATGCGCTTGTATACGAGCAAGCAATTAAGAGTGCATTAAGAGGCAATCGAG 549
547 GAATCGCTCACTACGATACGAGCAATGATGCTCATTAATTAACGCAAGCTGAATAT 606
550 CAATTCGCAATTCGCAATTAAGCAATGATGCTCATTAATTAACGCAAGCTGAATAT 609
607 AAAGGATGAATGATGATGCTCCGAGTATTTGCAAAAAGGGGCAATCCAGGCAATC 666
610 AAAGGATGAATGATGATGCTCCGAGTATTTGCAAAAAGGGGCAATCCAGGCAATC 669
667 GGTGTATGACAGAGCGCAATGCTTCCGAGTTCGCAATCTGATTAATGATTAACAGAGA 726
670 GGTGTATGACAGAGCGCAATGCTTCCGAGTTCGCAATCTGATTAATGATTAACAGAGA 729
727 AACGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
730 AATGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
787 CGGACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
790 CGGACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
847 GGAATGCGCGCAACGAATTAAGAGTGCACCTCAAGGAACTTGTCTTCAATCATC 906
850 GGAATGCGCGCTCACTAATTAAGAGTGCACCTCAAGGAACTTGTCTTCAATCATC 909
907 ATGATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
910 ATGATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
967 GCATTCAGTGAAGTGCAGAAATTCATCAAACTCTGAGGGGGGCGGCTGAATGGGGCC 1026
970 GCATTCAGTGAAGTGCAGAAATTCATCAAACTCTGAGGGGGGCGGCTGAATGGGGCC 1029
1027 TACACGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
1030 TACACGACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1089
1087 TTCGGGCTGGGAATGAAGCGGAGCGGCTACATCAGTGCACCTGCTGAGCTTAA 1146
1090 TTCGGGCTGGGAATGAAGCGGAGCGGCTACATCAGTGCACCTGCTGAGCTTAA 1149
1147 AACGCAATTAAGTGCAGAAATTCATCAAACTCTGAGGGGGGCGGCTGAATGGGGCC 1206
1150 AACGCAATTAAGTGCAGAAATTCATCAAACTCTGAGGGGGGCGGCTGAATGGGGCC 1209
1207 AATATTAACAGCTTGTGACAGATTTCTTCCGCTGAGCGGCAAAAGATGGGCGATCAAG 1266
1210 AATATTAACAGCTTGTGACAGATTTCTTCCGCTGAGCGGCAAAAGATGGGCGATCAAA 1269

1267 CCTGATGTATGCGCGCCAGGAGCATATCATTTTATCAGCAAGATCTTCTTGCACCCGAT 1326
1270 CCGATGTATGAGGACCGGGAACGTTCACTATCAGCAAGATCTTCTTGCACCCGAT 1329
1327 TCTCTCTTCTGAGGGAATCATGACAGCAAAATATGCTTATATGAGTGAACGTCATGCA 1386
1330 TCTCTCTTCTGAGGGAATCATGACAGCAAAATATGCTTATATGAGTGAACGTCATGCA 1389
1387 ACACCGATTTGTCGGGGAATGTTGCAACAGCTCCGTGAGCAATTTTGTGAAAAATAGAGA 1446
1390 ACACCGATTTGTCGGGGAATGTTGCAACAGCTCCGTGAGCAATTTTGTGAAAAATAGAGA 1449
1447 ATCACTCTTAAGCTTCTCTTATGAAAGCACTTGTGATGAGTGCCTGATGATGATGATGAT 1506
1450 ATCACTCTTAAGCTTCTCTTATGAAAGCACTTGTGATGAGTGCCTGATGATGATGATGAT 1509
1507 TTGGGTTATCCGAAACGGAACCAAGGATGAGGCGGAGTGAACCTGATTAATTCGTTGAC 1566
1510 TTGGGTTATCCGAAACGGAACCAAGGATGAGGCGGAGTGAACCTGATTAATTCGTTGAC 1569
1567 GTTGCCTATGTAAGCAATTCAGTGCCTATCACTAGCCAAAAAGCGACATATACCTTT 1626
1570 GTTGCCTATGTAAGCAATTCAGTGCCTATCACTAGCCAAAAAGCGACATATACCTTT 1629
1627 ACTGCAACGCGGCGCAAGCTTGAATAATCTCTGCTGATGATGATGATGATGATGATGATGAT 1686
1630 ACTGCTACTGCGGCAAGCTTGAATAATCTCTGCTGATGATGATGATGATGATGATGATGAT 1689
1687 ACTGCTACTGCTGTAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746
1680 ACAACTGCTTCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
1747 ACAAGATATGTCGGGAATGACTTCTCAGACACATTTGACATTAATGAGATGAGCGCAAT 1806
1750 ACAAGATATGTAAGGAATGACTTCTCAGACACATTTGACATTAATGAGATGAGCGCAAT 1809
1807 AACGTAAGAAATGATTTATTAATTCGCCCCCAAGTGAACATATACATGATGATGATGAT 1866
1810 AACGTAAGAAATGATTTATTAATGACACCAAAAGCGGGAAGTATTAATGATGATGATGAT 1869
1867 GCATATATATGCGGCTTGAACCAAACTTCTGCTTGGCAATTTGATGATGAT 1920
1870 GCTTATATGTAACCGGCTTGAACCAAACTTCTGCTTGGCAATTTGATGATGAT 1923

RESULT 9
US-10-820-714A-2
; Sequence 2, Application US/10820714A
; Publication No. US20050214922A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuyoshi
; APPLICANT: Kobayashi, Tohru
; APPLICANT: Sumitomo, Nobuyuki
; APPLICANT: Takimura, Yasushi
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251697US0
; CURRENT APPLICATION NUMBER: US/10/820,714A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106709
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1920)
; FEATURE:

NAME/KEY: 819_peptide
LOCATION: (1)...(618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619)...()
US-10-820-714A-2

Query Match 76.9%; Score 1477.2; DB 9; Length 1923;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

```
QY 7 AAGAGAGAGTGTGTTTATCTGTTTATCAGTCGACGCAATTCGTGACTGTGCTATTA 66
DB 10 AAGAGAGAGTGTGTTTATCTGTTTATCAGTCGACGCAATTCGTGACTGTGCTATTA 69
QY 67 AACATTCCTCGGCTGCTGATGACAGGACTTTGATCTGGAATTTTAAAGAAATTCAGCA 126
DB 70 AGTAATCCATCTGACGCTGCTGACAGGAAATTTGATCTGGAATTTTAAAGAAATTCAGCA 129
QY 127 ACAACGAGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186
DB 130 ACAACGAGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
QY 187 GAGTCGAGAAATGTGAACTTCTTAAAGGATGCTTAAAGAACTTGAACAGTACCGGCA 246
DB 190 GAGTCGAGAAATGTGAACTTCTTAAAGGATGCTTAAAGGATGCTTGAACAGTACCGGCA 249
QY 247 AATATATTAATCTCATATTTGCTCAATTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 306
DB 250 AATATATTAATCTCATATTTGCTCAATTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
QY 307 CTAGAGACAACTGAGGCAAGAAATTCCTGCACTATCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
DB 310 CTAGAGACAACTGAGGCAAGAAATTCCTGCACTATCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 367 TATGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
DB 370 TATGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
QY 427 TACTGCTGCAATTAACAAATAGATCCCAAGCTTTTCACAAAGGCGCATGACGCTGCTG 486
DB 430 TACTGCTGCAATTAACAAATAGATCCCAAGCTTTTCACAAAGGCGCATGACGCTGCTGCT 489
QY 487 AAAGGCTGCGCTGATGACGAGCAAGCAATTAAGAGTGCATTTAAGAGCATTCGAG 546
DB 490 AAAGGCTGCGCTGATGACGAGCAAGCAATTAAGAGTGCATTTAAGAGCATTCGAG 549
QY 547 GAATTCGCTCAGTACGTAAGCAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
DB 550 GAATTCGCTCAGTACGTAAGCAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
QY 607 AAGGTGATGATGATGCTGCGCAAGAGTATTTGCAAGCGGATGTCGACAGAGCATAC 666
DB 610 AAGGTGATGATGATGCTGCGCAAGAGTATTTGCAAGCGGATGTCGACAGAGCATAC 669
QY 667 GGTGTTGATGACAGGCGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
DB 670 GGTGTTGATGACAGGCGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 727 AACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
DB 730 AACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
QY 787 CGAGCGAATTAATGCAATGATACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
DB 790 CGAGCGAATTAATGCAATGATACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 849
QY 847 GGAATGCGCGCAAGCAATTAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
DB 850 GGAATGCGCGCAAGCAATTAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
QY 907 ATGATGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
```

```
DB 910 ATGATGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
QY 967 GCATTCAGTGCAGGTGCGCAAAATTCATCAAACTCTGCGGCGCGACGCGTGAATGCGGCC 1026
DB 970 GCATTCAGTGCAGGTGCGCAAAATTCATCAAACTCTGCGGCGCGACGCGTGAATGCGGCC 1029
QY 1027 TACAGCAAGATTCAGAAATGTGATGATGCTATGTAAGAAATTAATGATGATGCTGCT 1086
DB 1030 TACAGCAAGATTCAGAAATGTGATGATGCTATGTAAGAAATTAATGATGATGCTGCTGCT 1089
QY 1087 TTGCGGCGCTGAGAAATGAAGCGCGCAAGCGCGTACCATCATGCTGCTGCTGCTGCTGCT 1146
DB 1090 TTGCGGCGCGGAAATGAAGCGCGCAAGCGCGTACCATCATGCTGCTGCTGCTGCTGCTGCT 1149
QY 1147 AACGCAATTAACGTCGCGCAACCGAAACTGCGCTGCAAGCTTCTGCTGCTGCTGCTGCTGCT 1206
DB 1150 AATGCAATTAACGTCGCGCAACCGAAACTGCGCTGCAAGCTTCTGCTGCTGCTGCTGCTGCT 1209
QY 1207 AATATTAACGAGTGTGCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
DB 1210 AATATTAACGAGTGTGCAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269
QY 1267 CCTGATGCTATGCGCGCGCAAGCATATTTATGACAGAAATCTTCTGCTGCAACCGAT 1326
DB 1270 CCTGATGCTATGCGCGCGCAAGCATATTTATGACAGAAATCTTCTGCTGCAACCGAT 1329
QY 1327 TCTCTCTTCTGCGGCAATCATGACAGCAAAATATGCTATATGAGGTGAAACGTTCATGCGCA 1386
DB 1330 TCTCTCTTCTGCGGCAATCATGACAGCAAAATATGCTATATGAGGTGAAACGTTCATGCGCT 1389
QY 1387 ACAACGATGCTGCGGCAATGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
DB 1390 ACAACGATGCTGCGGCAATGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449
QY 1447 ATCACTCTTAAGCCTTCCCTATTTGAAGCAGCTTGTGATGCTGAGTGTGCTGCTGCTGCTGCT 1506
DB 1450 ATCACTCTTAAGCCTTCCCTATTTGAAGCAGCTTGTGATGCTGAGTGTGCTGCTGCTGCTGCT 1509
QY 1507 TTGGGTTATCCGAAACGAAACCAAGATGCGGCGGAGTGAACCTGGAATTAATGCTTGAAC 1566
DB 1510 TTGGGTTATCCGAAACGAAACCAAGATGCGGCGGAGTGAACCTGGAATTAATGCTTGAAC 1569
QY 1567 GTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1626
DB 1570 GTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629
QY 1627 ACTGCAACGCGCGGCAAGCAATTTGAAGTCTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1686
DB 1630 ACTGCAACGCGCGGCAAGCAATTTGAAGTCTCCCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1689
QY 1687 ACTACTGCTTCTGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
DB 1690 ACTACTGCTTCTGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749
QY 1747 ACAAGATATGCTGCGGAATGATCTTCTGACGACATTTTGAACATTAATCTGCGATGCGCGCAT 1806
DB 1750 ACAAGATATGCTGCGGAATGATCTTCTGACGACATTTTGAACATTAATCTGCGATGCGCGCAT 1809
QY 1807 AACGTAAGAAATGATATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866
DB 1810 AACGTAAGAAATGATATTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1869
QY 1867 GCATTAATATGCTGCGGTTGACCAAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 1870 GCATTAATATGCTGCGGTTGACCAAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
```

RESULT 10
US-10-820-712A-11
; Sequence 11, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION

APPLICANT: Okuda, Mitsuoshi
APPLICANT: Izawa, Yoshitomi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820,712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 11
LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus sp. KSM-kp9860
FEATURE:
NAME/KEY: CDS (1302)
LOCATION: (1)..(1302)
OTHER INFORMATION: protease KP9860
US-10-820-712A-11

Query Match 67.5%; Score 1295.6; DB 8; Length 1302;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1298; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 616 AATGATGCGCCAGAGATTGTCTCAAGCGGATGTGGCAAGACGCTACGGTTTGAT 675
DB 1 AATGATGCGCCAGAGATTGTCTCAAGCGGATGTGGCAAGACGCTACGGTTTGAT 60
QY 676 GACACAGGCGCAGATTGTGCGAGTTGCCGATCTGATGATGATGATGATGATGAT 735
DB 61 GACACAGGCGCAGATTGTGCGAGTTGCCGATCTGATGATGATGATGATGATGAT 120
QY 736 TCGATGATGAAAGCTTCCGCGGTAAATTAACAGACATATATGACTGGGTGCGAGAT 795
DB 121 TCGATGATGAAAGCTTCCGCGGTAAATTAACAGACATATATGACTGGGTGCGAGAT 180
QY 796 AATGGAATGATACGAACGGTATGTATCCCATGTGGCAGGTTCCGATATAGAAATGGC 855
DB 181 AATGGAATGATACGAACGGTATGTATCCCATGTGGCAGGTTCCGATATAGAAATGGC 240
QY 856 GCAACGAATTAAGAAATGGACCTCAAGCGAATCTGTTTTCATTCATCATGATGATAGC 915
DB 241 GCAACGAATTAAGAAATGGACCTCAAGCGAATCTGTTTTCATTCATCATGATGATAGC 300
QY 916 AGTGTGGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCAAACATTCAGT 975
DB 301 AGTGTGGGCTTGAAGGCTTGCCTTCAATCTGCAAACTTATTCAGCCAAACATTCAGT 360
QY 976 GCAAGTGCAGAAATCATACAAATCTCTGGGGGGGAGCGGTGAATGGGGCTTACAGCA 1035
DB 361 GCAAGTGCAGAAATCATACAAATCTCTGGGGGGGAGCGGTGAATGGGGCTTACAGCA 420
QY 1036 GATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095
DB 421 GATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1096 GGGAAATGAAGCGGAAACGGGGGTACATCAATGTCACCTGGTAAACGCCATA 1155
DB 481 GGGAAATGAAGCGGAAACGGGGGTACATCAATGTCACCTGGTAAACGCCATA 540
QY 1156 ACAGTGGCGGACCGGAAACCTGTGTCAGGCTTCGTTCCATGAGATTAATTTTAC 1215
DB 541 ACTGTGGCGGACCGGAAACCTGTGTCAGGCTTCGTTCCATGAGATTAATTTTAC 600
QY 1216 CACGTTGACAGATTCTTCCCGTGGCCCGCAAAAGATGGCGATCAAGCTGATGTC 1275
DB 601 CACGTTGACAGATTCTTCCCGTGGCCCGCAAAAGATGGCGATCAAGCTGATGTC 660
QY 1276 ATGGCGCAGGAGACATATTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCTTC 1335
DB 1276 ATGGCGCAGGAGACATATTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCTTC 1335

DB 661 ATGGCGCAGGAGACATATTTATCAGCAAGATCTTCTTTCGACCCGATTCCTCTTC 720
QY 1336 TGGCGGAATCATGACAGAAATATGCTTATATGGGTGGAAGCTCAATGGCAACCGATT 1395
DB 721 TGGCGGAATCATGACAGAAATATGCTTATATGGGTGGAAGCTCAATGGCAACCGATT 780
QY 1396 GTTGCGGGGAATGTGACAGCTCCGTGAGATTTTGGAAAAATAGAGAAATCACTCT 1455
DB 781 GTTGCGGGGAATGTGACAGCTCCGTGAGATTTTGGAAAAATAGAGAAATCACTCT 840
QY 1456 AAGCTTCCCTATTGAAAGCAGCTTTGATGTCAGGTGCTGATGATGATGATGATGAT 1515
DB 841 AAGCTTCCCTATTGAAAGCAGCTTTGATGTCAGGTGCTGATGATGATGATGATGAT 900
QY 1516 CCGAACGGAAACCAAGATGGGGCCGAGTGAACCTCGATTAATGTTGAAGCTTGCTAT 1575
DB 901 CCGAACGGAAACCAAGATGGGGCCGAGTGAACCTCGATTAATGTTGAAGCTTGCTAT 960
QY 1576 GTGAACGAATCCAGTGCCTTATCAACTAGCCAAAGGACATATACCTTATGCAACG 1635
DB 961 GTGAACGAATCCAGTGCCTTATCAACTAGCCAAAGGACATATACCTTATGCAACG 1020
QY 1636 GCGGGCAAGCATTGAAATCTCCCTGTGATGTCGATGTCCTGCAAGCACTACTGCT 1695
DB 1021 GCGGGCAAGCATTGAAATCTCCCTGTGATGTCGATGTCCTGCAAGCACTACTGCT 1080
QY 1696 TCTGTAACCTGTGCAATGATTTGATTTGCTATACAGACCAAAACGAAACGATAT 1755
DB 1081 TCTGTAACCTGTGCAATGATTTGATTTGCTATACAGACCAAAACGAAACGATAT 1140
QY 1756 GTCGGAATGACTTCTGACGACATTTGACATTAATGAGATGAGCCCAATTAAGTAA 1815
DB 1141 GTCGGAATGACTTCTGACGACATTTGACATTAATGAGATGAGCCCAATTAAGTAA 1200
QY 1816 AATGTAATTAATTAATTCGCCCAAGTGAAACATATACATTAAGTGTGCAATATAT 1875
DB 1201 AATGTAATTAATTAATTCGCCCAAGTGAAACATATACATTAAGTGTGCAATATAT 1260
QY 1876 GTGCGGTTGACCAAAACCTTCTGTTGGCAATTTGTGAAC 1917
DB 1261 GTGCGGTTGACCAAAACCTTCTGTTGGCAATTTGTGAAC 1302

RESULT 11

US-10-820-714A-12

; Sequence 12, Application US/10820714A

; Publication No. US20050214922A1

; GENERAL INFORMATION:

; APPLICANT: KAO CORPORATION

; APPLICANT: Okuda, Mitsuoshi

; APPLICANT: Kobayashi, Tohru

; APPLICANT: Sumitomo, Nobuyuki

; APPLICANT: Takimura, Yasushi

; TITLE OF INVENTION: ALKALINE PROTEASE

; FILE REFERENCE: 251697USO

; CURRENT APPLICATION NUMBER: US/10/820,714A

; PRIOR FILING DATE: 2004-04-09

; PRIOR APPLICATION NUMBER: 2003-106709

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 12

; LENGTH: 1302

; TYPE: DNA

; ORGANISM: Bacillus sp. KSM-kp9860

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1302)

; OTHER INFORMATION: protease KP9860

US-10-820-714A-12

Query Match

67.5%; Score 1295.6; DB 9; Length 1302;

Best Local Similarity 99.7%; Pred. No. 0;
Matches 1298; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 616 AATGATGTGGCCGAGGTATTTTCAGAAAGCGGATGTGGCAGAGCACTACGGTTGTAT 675
DB 1 AATGATGTGGCCGAGGTATTTTCAGAAAGCGGATGTGGCAGAGCACTACGGTTGTAT 60
QY 676 GGAAGAAGCCAGATTGTGGCAGTGTGGCAATTCGATTCAGAGAAAGAAAGCAAGT 735
DB 61 GGAAGAAGCCAGATTGTGGCAGTGTGGCAATTCGATTCAGAGAAAGAAAGCAAGT 120
QY 736 TCGATGATGAAGCTTCCGCGGTAAATTAACAGCACTATATGCTAGGTCCGAGAAAT 795
DB 121 TCGATGATGAAGCTTCCGCGGTAAATTAACAGCACTATATGCTAGGTCCGAGAAAT 180
QY 796 AATGCGAATGATTCGAAAGCGTATGTGATCCATGTGGCAGGTTCGGTATTAGAAATGCG 855
DB 181 AATGCGAATGATTCGAAAGCGTATGTGATCCATGTGGCAGGTTCGGTATTAGAAATGCG 240
QY 856 GCAACGAATTAAGGAATGGCACTCAAGGGAATCTGGTTTTCAATCCATCATGATAGC 915
DB 241 GCAACGAATTAAGGAATGGCACTCAAGGGAATCTGGTTTTCAATCCATCATGATAGC 300
QY 916 AGTGTGGCTTGGAGCTTGCCTTCCAAATCTGCAAACTTATTCAGCCAGCAATTCAGT 975
DB 301 AGTGTGGCTTGGAGCTTGCCTTCCAAATCTGCAAACTTATTCAGCCAGCAATTCAGT 360
QY 976 GCAAGTGGCAATTCATACAACTCTCTGGGGGGGACGGGTGATGGGGCTTACAGACA 1035
DB 361 GCAAGTGGCAATTCATACAACTCTCTGGGGGGGACGGGTGATGGGGCTTACAGACA 420
QY 1036 GATTCGAGAAATGTGATGATCTATGTAAGAAATGATATGACGATCTTTTTCGGGCT 1095
DB 421 GATTCGAGAAATGTGATGATCTATGTAAGAAATGATATGACGATCTTTTTCGGGCT 480
QY 1096 GGAATGAAGGCGCAACGCGGTACCATCATGTCACCTGTGAGGCTAAAAAGCCATA 1155
DB 481 GGAATGAAGGCGCAACGCGGTACCATCATGTCACCTGTGAGGCTAAAAAGCCATA 540
QY 1156 ACAGTGGCGCAACCGAAAACTGCGTCCAGCTTCGGTTCTATGCAATATATTAAAC 1215
DB 541 ACAGTGGCGCAACCGAAAACTGCGTCCAGCTTCGGTTCTATGCAATATATTAAAC 600
QY 1216 CAGTTGCAAGTTCTTCCGTGGCCGCAAGAAAGATGGGGATTCAGGCTGATGTC 1275
DB 601 CAGTTGCAAGTTCTTCCGTGGCCGCAAGAAAGATGGGGATTCAGGCTGATGTC 660
QY 1276 ATGGCGCCAGGACATACATTTTATCAGAAAGATCTTCTTGGCAACCGATTCCTCTTC 1335
DB 661 ATGGCGCCAGGACATACATTTTATCAGAAAGATCTTCTTGGCAACCGATTCCTCTTC 720
QY 1336 TGGGCGAATCATGACAGCAAAATATGCTATATGGGTGAACGTCATGCGCAACCGATT 1395
DB 721 TGGGCGAATCATGACAGCAAAATATGCTATATGGGTGAACGTCATGCGCAACCGATT 780
QY 1396 GTTGGCGGAATGTTCACAGCTCCGTGAGCATTTTGTGAAAAATAGAGAAATCACTCT 1455
DB 781 GTTGGCGGAATGTTCACAGCTCCGTGAGCATTTTGTGAAAAATAGAGAAATCACTCT 840
QY 1456 AAGCCTTCCCTATTGAAAGCAGCTTGTGATGGAGGTGCTGATGTTGATTTGGTTAT 1515
DB 841 AAGCCTTCCCTATTGAAAGCAGCTTGTGATGGAGGTGCTGATGTTGATTTGGTTAT 900
QY 1516 CCGAAGCGAAGCAAGATGGGGCGAGTGACCTTGATTAATTCGTTGAACGTTGCTAT 1575
DB 901 CCGAAGCGAAGCAAGATGGGGCGAGTGACCTTGATTAATTCGTTGAACGTTGCTAT 960
QY 1576 GTGAAGCAATTCAGTGGCTTATCACTACCAAAAAAGCAATATACCTTTTACTGGAAG 1635
DB 961 GTGAAGCAATTCAGTGGCTTATCACTACCAAAAAAGCAATATACCTTTTACTGGAAG 1020
QY 1636 GGGGCAAGCATTGAATTCCTCGTGTATGATGGGATGGCCCTGCAAGCACTACTGCT 1695
DB 1695 GGGGCAAGCATTGAATTCCTCGTGTATGATGGGATGGCCCTGCAAGCACTACTGCT 1695

DB 1021 GGGGCAAGCATTGAATTCCTCGTGTATGATGGGATGGCCCTGCAAGCACTACTGCT 1080
QY 1696 TCTGTAACTCTGTCAATGATTTTGGATTTGTGATTCACGACCAAAAGCAATAT 1755
DB 1081 TCTGTAACTCTGTCAATGATTTTGGATTTGTGATTCACGACCAAAAGCAATAT 1140
QY 1756 GTCCGGAATGATCTTCCAGCAACATTTGCAATTAACGTGGAGGCGCAATACGTAAGA 1815
DB 1141 GTCCGGAATGATCTTCCAGCAACATTTGCAATTAACGTGGAGGCGCAATACGTAAGA 1200
QY 1816 AATGATTTTATTAATTCGCCCAAGGTGAAACATATACCATTTAGGTGCAAGCATTAAT 1875
DB 1201 AATGATTTTATTAATTCGCCCAAGGTGAAACATATACCATTTAGGTGCAAGCATTAAT 1260
QY 1876 GTCCGCTTGAACCAAAACTTCTCGTTGGCAATTTGTAAC 1917
DB 1261 GTCCGCTTGAACCAAAACTTCTCGTTGGCAATTTGTAAC 1302

RESULT 12
US-10-385-662-1
; Sequence 1, Application US/10385662
; Publication No. US20040002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: SAITO, KAZUHIRO
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIYUKI
; APPLICANT: SAKI, KATSUHIISA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease
; FILE REFERENCE: 2349380S0
; CURRENT APPLICATION NUMBER: US/10/385,662
; CURRENT FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: JP 2002-081428
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP43
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1305)
; OTHER INFORMATION:
US-10-385-662-1

Query Match 52.2%; Score 1002.6; DB 6; Length 1305;
Best Local Similarity 85.5%; Pred. No. 6; 6e-299;
Matches 1116; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY 616 AATGATGTGGCCGAGGTATTTTCAGAAAGCGGATGTGGCAGAGCACTACGGTTGTAT 675
DB 1 AATGATGTGGCCGAGGTATTTTCAGAAAGCGGATGTGGCAGAGCACTACGGTTGTAT 60
QY 676 GGAAGAAGCCAGATTGTGGCAGTGTGGCAATTCGATTCAGAGAAAGAAAGCAAGT 735
DB 61 GGAAGAAGCCAGATTGTGGCAGTGTGGCAATTCGATTCAGAGAAAGAAAGCAAGT 120
QY 736 TCGATGATGAAGCTTCCGCGGTAAATTAACAGCACTATATGCTAGGTCCGAGAAAT 795
DB 121 TCGATGATGAAGCTTCCGCGGTAAATTAACAGCACTATATGCTAGGTCCGAGAAAT 180
QY 796 AATGCGAATGATTCGAAAGCGTATGTGATCCATGTGGCAGGTTCGGTATTAGAAATGCG 855
DB 181 AATGCGAATGATTCGAAAGCGTATGTGATCCATGTGGCAGGTTCGGTATTAGAAATGCG 855

Db 1261 GTTCCACAGGGCCGCGTTTCTTGGCTATTGTGAC 1302

Search completed: April 8, 2006, 07:34:32
Job time : 1381.42 secs

14

This Page Blank (usp10)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 03:01:47 ; Search time 896.3 Seconds
(without alignments)

8569.753 Million cell updates/sec

Title: US-10-784-870-3

Perfect score: 1920

Sequence: 1 atgagaagaagaagtgctt.....cgttcgaatcgtgaactaa 1920

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New.*

Result No.	Score	Query Match	Length	ID	Description
1	37	1.9	972	US-10-301-480-95183	Sequence 95183, A
2	37	1.9	972	US-10-301-480-708592	Sequence 708592, A
3	36.6	1.9	622	US-09-925-065A-613029	Sequence 613029, A
4	36.2	1.9	600	US-10-301-480-70889	Sequence 70889, A
5	36.2	1.9	628	US-10-301-480-692098	Sequence 692098, A
6	36	1.9	628	US-09-925-065A-600872	Sequence 600872, A
7	36	1.9	628	US-09-925-065A-600873	Sequence 600873, A
8	36	1.9	172649	US-11-121-086-36	Sequence 36, App1
9	35.8	1.9	549	US-09-925-065A-609963	Sequence 609963, A
10	35.8	1.9	580	US-10-301-480-70454	Sequence 70454, A
11	35.8	1.9	580	US-10-301-480-70455	Sequence 70455, A
12	35.8	1.9	580	US-10-301-480-683863	Sequence 683863, A
13	35.8	1.9	580	US-10-301-480-683865	Sequence 683865, A
14	35.8	1.9	650	US-09-925-065A-822653	Sequence 822653, A
15	35.6	1.9	431	US-09-925-065A-600874	Sequence 600874, A
16	35.6	1.9	753	US-10-932-182A-4129	Sequence 4129, App
17	35.6	1.9	753	US-10-932-182A-4129	Sequence 4129, App
18	35.6	1.9	4656	US-10-932-182A-3130	Sequence 3130, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

19	35.6	1.9	4656	US-10-932-182A-3130	Sequence 3130, App
20	35.4	1.8	580	US-10-301-480-70455	Sequence 70455, A
21	35.4	1.8	580	US-10-301-480-683864	Sequence 683864, A
22	35.4	1.8	1175	US-11-102-026A-111	Sequence 111, App
23	35.4	1.8	1215	US-11-102-026A-187	Sequence 187, App
24	35	1.8	1758	US-10-821-234-637	Sequence 637, App
25	35	1.8	611587	US-11-117-187-209	Sequence 209, App
26	34.6	1.8	390183	US-10-330-773-937	Sequence 937, App
27	34.4	1.8	709	US-09-925-065A-866816	Sequence 866816, A
28	34.4	1.8	995	US-10-301-480-564083	Sequence 564083, A
29	34.4	1.8	995	US-10-301-480-1177492	Sequence 1177492, A
30	34.4	1.8	159146	US-11-121-086-49	Sequence 49, App1
31	34.2	1.8	452	US-09-925-065A-199670	Sequence 199670, A
32	34.2	1.8	455	US-10-301-480-288350	Sequence 288350, A
33	34.2	1.8	455	US-10-301-480-901759	Sequence 901759, A
34	34.2	1.8	5617	US-11-136-527-2014	Sequence 2014, App
35	34	1.8	515	US-09-925-065A-701445	Sequence 701445, A
36	34	1.8	722	US-09-925-065A-85259	Sequence 85259, A
37	34	1.8	722	US-09-925-065A-85260	Sequence 85260, A
38	34	1.8	722	US-10-301-480-186499	Sequence 186499, A
39	34	1.8	722	US-10-301-480-186500	Sequence 186500, A
40	34	1.8	722	US-10-301-480-799908	Sequence 799908, A
41	34	1.8	722	US-10-301-480-799909	Sequence 799909, A
42	34	1.8	998	US-10-301-480-562054	Sequence 562054, A
43	34	1.8	998	US-10-301-480-1175463	Sequence 1175463, A
44	34	1.8	1051	US-09-925-065A-64075	Sequence 64075, A
45	34	1.8	1051	US-10-301-480-165313	Sequence 165313, A

ALIGNMENTS

RESULT 1
US-10-301-480-95183
Sequence 95183, Application US/10301480
Publication No. US2006057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 95183
LENGTH: 972
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-95183

Query Match 1.9%; Score 37; DB 9; Length 972;
Best Local Similarity 51.5%; Pred. No. 4.4;
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

1167 AACGAAACCTGCTCCAGGTTGCTTCATGACATATTAACCAAGTTCGACA 1226
525 AATGCTCAAGGCACTTAACTTAAGTTCATTAATTAATCACTTAAGAACTTGA 584
1227 GTTCTTCCTCCGTCGACCAAAAGATGGGAGATCAAGCTGATGTCATGCGCCAG 1286
555 ATTCTGATGCCAGCCACATCATTAATTAATTAATCTCTTAAGAGAGATCCAGA 644
1287 GACATACATTTTATCAGCAAGATCTTCTTTCGACCCGATTCCTC 1331
645 CATGATATGTTTCTTAAGTCTCCCTAACCCAGGATTCAC 689

RESULT 2

```
US-10-301-480-708592
; Sequence 708592, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1225818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 708592
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-708592
```

```
Query Match          1.9%; Score 37; DB 10; Length 972;
Best Local Similarity 51.5%; Pred. No. 4.4;
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
```

```
Oy 1167 AACGAAACCTGCTCCAGCTTGGTTCCTATGCAGATATATTAACCAAGTTGCACA 1226
Db 525 AATGTCAGAGGCGCATTTAACTTAGTGCAATTTAAAGATCACTTAGAAGACTTTGAAA 584
Oy 1227 GTTCTCTCCGCGGCGGCAAAAGATGGCGAATCAAGCTGATGTCATGGCGCCAG 1286
Db 585 ATTCTGAGCCGAGCCGACATCATATGAAATTAACCTCTAAGAGGAGAGATCCAGA 644
Oy 1287 GACATACATTTTATCAGCAAGATCTTCTCTGACACCCCAATTCCTC 1331
Db 645 CATTGATAGTTTCTTAAGTGTCCCTACCCCAAGGTGATTCAC 689
```

```
RESULT 3
US-09-925-065A-613029/c
; Sequence 613029, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 613029
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-613029
```

```
Query Match          1.9%; Score 36.6; DB 6; Length 622;
Best Local Similarity 46.6%; Pred. No. 4.6;
Matches 117; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
```

```
Oy 109 TTTAAGGAATTCAGCAAGCAACGATGTCAGTGTCTTCTCCAAACAGCGAACAAGGT 168
```

```
Db 441 TTTTTCGAATCABAATTTTGAATATGTATGAAAGCTTTAAACATCCATTAATCAAT 382
Oy 169 GCGGCTGATTTCTGCGAGACTGGAANAATGTAACCTTTTAAAGATTCTTAAGAAA 228
Db 381 CATTGAGATTTTCACTTTTGGTCCCAACATTTACATACATAGATTTGTGTAAAT 322
Oy 229 CTGAACAGTACCGGCAATTAATTAACCTCATATTGTCCAATTCAATGCCCATTTTA 288
Db 321 GCAACTAATGTATTCCAAATTAAGTTTCTTATGTGAGATTTAGCAAGAGCTCTTGA 262
Oy 289 GAAAGAAACAAACAGAGCTAGAGACAACAGCAAGATTTCTGACTACATCCCTGAT 348
Db 261 ATTCTTCAAAATGAGAACTTAATATATATGCAATATATATATATATTAATTTTAT 202
Oy 349 TATGCAATATAT 359
Db 201 AATATGTATAT 191
```

```
RESULT 4
US-10-301-480-78689/c
; Sequence 78689, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78689
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-78689
```

```
Query Match          1.9%; Score 36.2; DB 9; Length 600;
Best Local Similarity 54.2%; Pred. No. 5.9;
Matches 71; Conservative 1; Mismatches 59; Indels 0; Gaps 0;
```

```
Oy 177 ATTTCTGAGAGTCTGAATAATGTGAACCTTTAAAGATTGCTAAAGAACTTGAAAC 236
Db 420 ATATAGGGGAAAGATTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 361
Oy 237 AGTACCGCAATATTAATTAACCTCATATTGTCATTAATGAGCCCATTTTAAAGAAAC 296
Db 360 TTAACAGGCAAAATATATGACAGAAAAAATTAATTAATTAATTAATTAATTAATTA 301
Oy 297 AAAACAGAGC 307
Db 300 ACTTAAAAATC 290
```

```
RESULT 5
US-10-301-480-692098/c
; Sequence 692098, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
```


Query Match	1.9%	Score 36;	DB 14;	Length 172649;
Best Local Similarity	67.1%;	Pred. No. 1.5e+02;		
Matches 51;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0
0y	192	TGAATATGTGAACCTTCTTAAGAATTCCTAAGAACTGAAACAGTACCGGCAATTA	251	

Db 141108 TGAAGATGGCACTTATATGTGAGATGTAAGAAAGTTGAACTCATGAGACGAGAA 141167

Qy 252 TAAACTCCATATTTGTC 267

Db 141168 TAAATGGTTATTC 141183

RESULT 9

US-09-925-065A-609963/c
; Sequence 609963, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 609963
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-609963

Query Match 1.9%; Score 35.8; DB 6; Length 549;

Best Local Similarity 63.2%; Pred. No. 7.4; Matches 55; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 194 AAAATGTGAACCTCTTAAGGATTCCTAAAGAACTGAAACAGTACCGCAATATATA 253

Db 87 AAAATGAGAGATTTTAAATATATAGCCACAAAATATGATTAAGAAAAATTT 28

Qy 254 AACTCCATATTTGCAATTCATGACC 280

Db 27 AAGATCTTTTAAATCTTTTCC 1

RESULT 10

US-10-301-480-70454/c
; Sequence 70454, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70454
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-70454

Query Match 1.9%; Score 35.8; DB 9; Length 580;

Best Local Similarity 48.7%; Pred. No. 7.7; Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 634 ATTGTCAAGCGGATGTGGACAGACGACTACGTTTGTATGCAAGCCAGATTGTC 693

Db 374 ATGTTCTTGAGAGCTGAGACGACTAGTTACTTGGGGTATAACATGGGGTCCGCTG 315

Qy 694 GCATTCGCCATCTGATTTGATATACAGAAAGAACGACTTCGATGCATGAAGCCCTTC 753

Db 314 GCACATGCCAGCAATTTAGACACAGAAACACACAGAGATTTTAGAGTGAAGTTTA 255

Qy 754 CGCGTAAATATACAGCACTATATGCACTGGGTGCGAGCAATTAATGCAATGATACGAA 813

Db 254 ATGGGTAGAGAGAGAGAGAAAGAAACTGCTTCTTAAAGAGAAAGGGCTTTGAAAC 195

Qy 814 GGTCAATGATCCCATGTGG 832

Db 194 GGAAGAAAGACCATCTGG 176

RESULT 11

US-10-301-480-70456/c
; Sequence 70456, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70456
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-70456

Query Match 1.9%; Score 35.8; DB 9; Length 580;

Best Local Similarity 48.7%; Pred. No. 7.7; Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 634 ATTGTCAAGCGGATGTGGACAGACGACTACGTTTGTATGCAAGCCAGATTGTC 693

Db 374 ATGTTCTTGAGAGCTGAGACGACTAGTTACTTGGGGTATAACATGGGGTCCGCTG 315

Qy 694 GCATTCGCCATCTGATTTGATATACAGAAAGAACGACTTCGATGCATGAAGCCCTTC 753

Db 314 GCACATGCCAGCAATTTAGACACAGAAACACACAGAGATTTTAGAGTGAAGTTTA 255

Qy 754 CGCGTAAATATACAGCACTATATGCACTGGGTGCGAGCAATTAATGCAATGATACGAA 813

Db 254 ATGGGTAGAGAGAGAGAGAAAGAAACTGCTTCTTAAAGAGAAAGGGCTTTGAAAC 195

Qy 814 GGTCAATGATCCCATGTGG 832

Db 194 GGAAGAAAGACCATCTGG 176

RESULT 12

US-10-301-480-683863/c
; Sequence 683863, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.137

```
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 683863
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-683863
```

```
Query Match      1.9%; Score 35.8; DB 10; Length 580;
Best Local Similarity 48.7%; Pred. No. 7.7;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
```

```
Qy 634 ATTGTCAAAGCGATGTGGCAGACGACTACGGTTTGTATGACAAAGCCAGATTGTC 693
Db 374 ATGTTCTTGAGAGAGCTGAAGCAGCTAGTTACTTGGGGTATTAACATGGGCTTCGCTG 315
Qy 694 GCAAGTTCGCGATCTGATGTGATACAGAGAAAGACAGTTCGATGATGAAGCTTC 753
Db 314 GCACATGCCAGACATTTAGGACACAGAAACACAGAGATTTTAGAGTGGAGTTTA 255
Qy 754 CGCGGTAAATTAACAGCACTATATGACTGGGTGGACGATTAATGCCAATGATGCAAC 813
Db 254 ATGGGTAGAGAGAAAGAGAAAGAACTGCTTCTCTAAAGAGAGGGGCTTGGAAAC 195
Qy 814 GGTCAATGTAACCATGTGG 832
Db 194 GGAAGAAAAAGACATCTGG 176
```

```
RESULT 13
US-10-301-480-683865/c
/ Sequence 683865, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827,137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 683865
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-683865
```

```
Query Match      1.9%; Score 35.8; DB 10; Length 580;
Best Local Similarity 48.7%; Pred. No. 7.7;
Matches 97; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
```

```
Qy 634 ATTGTCAAAGCGATGTGGCAGACGACTACGGTTTGTATGACAAAGCCAGATTGTC 693
Db 374 ATGTTCTTGAGAGAGCTGAAGCAGCTAGTTACTTGGGGTATTAACATGGGCTTCGCTG 315
Qy 694 GCAAGTTCGCGATCTGATGTGATACAGAGAAAGACAGTTCGATGATGAAGCTTC 753
Db 314 GCACATGCCAGACATTTAGGACACAGAAACACAGAGATTTTAGAGTGGAGTTTA 255
Qy 754 CGCGGTAAATTAACAGCACTATATGACTGGGTGGACGATTAATGCCAATGATGCAAC 813
Db 254 ATGGGTAGAGAGAGAAAGAGAAAGAACTGCTTCTCTAAAGAGAGGGGCTTGGAAAC 195
```

```
Qy 814 GGTCAATGTAACCATGTGG 832
Db 194 GGAAGAAAAAGACATCTGG 176
```

```
RESULT 14
US-09-925-065A-822653
/ Sequence 822653, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 822653
/ LENGTH: 650
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-822653
```

```
Query Match      1.9%; Score 35.8; DB 6; Length 650;
Best Local Similarity 57.7%; Pred. No. 8.2;
Matches 64; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
Qy 1745 GAACAAGATATGTGCGAATGACTTCTGACGACCAATTGACAAATTAATGCGATGCGCGCA 1804
Db 146 GAATATGATATTTAGTGAATGTTTCATATACCAACAAACAAACCTATCATTAATCTGCA 205
Qy 1805 ATAAGTGAAGAAATGATTTTATTAATTCGCCCAAGTGAACATATACCA 1855
Db 206 CAATATTAAGAAATTTATTTGCAATTTATCCAAATTTTAAGCATACCA 256
```

```
RESULT 15
US-09-925-065A-600874
/ Sequence 600874, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 600874
/ LENGTH: 431
```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-600874

Query Match 1.9%; Score 35.6; DB 6; Length 431;
Best Local Similarity 54.8%; Pred. No. 7.5;
Matches 68; Conservative 1; Mismatches 55; Indels 0; Gaps 0;

QY	203	AACTCTTAAGGATTGCTAAAGAACTTGAACAGTACCGGCAATATTAATCACTCATA	262
DB	109	AACATTCTACAAAGTTCTTAAAGTCTCAGCAGGCTTGAGTTGTAATTGACCAATT	168
QY	263	TTGTCGAATTCATGCGCCCATTTTGAAGAAACAAACAGAGCTAGAGCAACTGAG	322
DB	169	TTAACCACTCATGACCTTTTWTWAAAAAAAAAAAAAAAAAAGACTGTGGCTA	228
QY	323	CAAA	326
DB	229	CAAA	232

Search completed: April 8, 2006, 03:55:10
Job time : 898.3 secs

Y

Y

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search **us-10-784-870-4.p2n.rge**

Run on: Apr 11 8, 2006, 03:02:08 ; Search time 7456 Seconds
(without alignments)
4871.642 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267

Sequence: 1 MRKKVFLSVLSAAALSTV.....EVQAVNPVGPNFSLATVN 639

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=/abs/ABSSWB.epool/US10784870/runat_07042006_143415_27459/app_query.fasta_1
-DB=GenEmbl -QMT=fastrap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=ad8803p
-USER=US10784870 @CGN_1_1_7415 @runat_07042006_143415_27459 -NCPU=6 -ICPU=3
-NO MMAP -MSG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_da:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_str:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3267	100.0	1920	1	AB046403 Bacillus
2	3267	100.0	1920	6	AR368116 Sequence
3	3267	100.0	1920	6	AR562464 Sequence

4	3072	94.0	1923	6	AR368118 Sequence
5	3072	94.0	1923	6	AR562466 Sequence
6	3066	93.8	1923	1	AB051423 Bacillus
7	3066	93.8	1923	6	CQ888285 Sequence
8	3066	93.8	1923	6	CQ889309 Sequence
9	3066	93.8	1923	6	AR368117 Sequence
10	3066	93.8	1923	6	AR562465 Sequence
11	3053	93.4	1923	1	AB084155 Bacillus
12	2799	85.7	3003	6	AR069954 Sequence
13	2799	85.7	3003	6	BD062155 Sequence
14	2500.5	76.5	2218	6	E03808 DNA encodin
15	2183	66.8	1305	6	AX839476 Sequence
16	2122	65.0	1302	1	AB046406 Bacillus
17	2002.5	61.3	1299	1	AB046405 Bacillus
18	1998.5	61.2	1299	1	AB046402 Bacillus
19	1991.5	61.0	1299	1	AB046404 Bacillus
20	816	25.0	346148	1	BX842652 Bdellovib
21	680	20.8	60006	3	AF268611 Unculture
22	655	20.0	110000	1	AE017180_35
23	566	17.3	5890	2	DD060086_
24	541	16.6	5772	2	AF466309 Dictyoste
25	516.5	15.8	6115	2	DD020432 Dictyoste
26	511	15.6	5259	2	AF263455 Dictyoste
27	477	14.6	1977	6	AR201152 Sequence
28	477	14.6	1977	6	AR578779 Sequence
29	477	14.6	1977	6	AR633917 Sequence
30	443	13.6	1962	6	AR201155 Sequence
31	443	13.6	1962	6	AR578782 Sequence
32	443	13.6	1962	6	AR633946 Sequence
33	443	13.6	9125	6	AR619564 Sequence
34	443	13.6	12452	1	AE010265 Pyrococcu
35	435	13.3	14376	1	AE013049 Thermococ
36	435	13.3	110000	1	AF006878_14
37	432	13.2	1977	6	AR633919 Sequence
38	431	13.2	14351	2	AC096673 Trypanos
39	417.5	12.8	1236	6	AR201146 Sequence
40	417.5	12.8	1236	6	AR578773 Sequence
41	417.5	12.8	1566	6	AR633918 Sequence
42	402.5	12.3	303450	1	SC093910 Scryptomy
43	401.5	12.3	110000	1	BA000028_24
44	376	11.5	110000	1	CP000099_19
45	376	11.5	110000	1	CP000099_20

ALIGNMENTS

RESULT 1	AB046403	1920 bp	DNA	linear	BCT 10-MAY-2002
LOCUS	Bacillus sp. 9860	PROA gene for protease, complete cds.			
DEFINITION	AB046403				
ACCESSION	AB046403.2	GI:20521152			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Bacillus sp. 9860				
REFERENCE	Bacillus sp. 9860				
AUTHORS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
TITLE	1				
	Saeki, K., Okuda, M., Hatada, Y., Kobayashi, T., Ito, S., Takami, H. and				
	Horikoshi, K.				
	Novel oxidatively stable subtilisin-like serine proteases from				
	alkaliphilic Bacillus spp.: enzymatic properties, sequences, and				
	evolutionary relationships				
JOURNAL	Biochem. Biophys. Res. Commun. 279 (2), 313-319 (2000)				
PUBMED	11118284				
REFERENCE	2 (bases 1 to 1920)				
AUTHORS	Saeki, K.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2000) Katsuhisa Saeki, Kao Corporation,				
	Biological Science Laboratory; Ichikaimachi Akabane 2606, Hoga,				
	Tochigi 321-3497, Japan (E-mail:387185@kasanet.kao.co.jp,				
	Tel:81-285-68-7400, Fax:81-285-68-7403)				
COMMENT	On May 9, 2002 this sequence version replaced gi:12381938.				
FEATURES	Location/Qualifiers				

```
source
1. 1920
/organism="Bacillus sp. 9860"
/mol_type="genomic DNA"
/strain="9860"
/db_xref="taxon:133778"
1. 1920
/gene="PROA"
1. 1920
/gene="PROA"
/codon_start=1
/cranal_table=1
/product="protease"
/protein_id="BAB21266.2"
/db_xref="gi:20521153"
/cranal_table=1
/translation="MRKKVFLSVLSAAILSTVALNNSAGDARTFDLPKGIQTTF
DVGSPSKOROTGAALFLVESBNVKILKGLKLETPVANNKLIHVOFNPLIBETKOK
LETTAKLIDYI PDYAYIYEYGDQSVKRSI EHVESYEPYLPKIDPDLTGAST
LVKALADITKONNKEVOLRGEIEIAQVYASNDVHYITAKPEYKMNNDVARGIVKADVA
OSYGLVGGQGIIVADTGLDPTGDSMHEAFRGKITALVALGRTNANDTNGHTH
VAGSVLNGATRKGAPOANLVPQSIHDSGGLGSPNLDTLPSQAFSAGARLHTNS
WGAAVNGAYTTDSNRVDYVRKNDMTILFAAGNERPNGTISAPGTAKNALIVATEN
LRPFGYADNIHVADPSSRGPTDGRKIDVMAPGTYIISASSIAPDSFVAMND
SKYAMGTSMATPIVAGNVAQLREHFVKNRGITPKPSILKALILAGAADGLGPN
VNOGMRVTLDKSLNVAYNVESALSTSOKATYTFATGKPLSLVMSDAPASTAS
VTLVNDLIVITAPNGRTYVGNDSAPFDNMDCRNVENVFINSPOGTYTIEQAY
NPVGPQNFSLATYN"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 3,14e-169 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0
```

US-10-784-870-4 (1-639) x AB046403 (1-1920)

```
QY 1 MetArglyValValPheLeuSerValLeuSerAlaAlaAlaLeuSerThrVal 20
Db 1 ATGAGAAAGAAAGAGGTCTTTTATCTGTTTATCATCGACGCAATTCCTCGACTGTT 60
QY 21 AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheGlyLe 40
Db 61 GCATTAAACAATCCCTCGGCTGGTGAATGACAGACATTTTGATCTGGATTTTAAAGAAAT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerIysGlnArgGlnThrGlyAlaAlaPhe 60
Db 121 CAACCAACAACCGATGTCAGTGGTTTCTCCAAACAGCACAACAAGTGGCGGTCATTT 180
QY 61 LeuValGluSerGluAsnValIleLeuLeuIleGlyLeuLeuValValLeuGluThrVal 80
Db 181 CTGGTGGAGTCTGAAATGTAACCTTCTTAAAGATTCCTAAAGAACTTGAACAGTA 240
QY 81 ProAlaAsnAsnIleAsnIleValGlnPheAsnGlyProIleLeuGluGluThrIys 100
Db 241 CCGGAAATATTAATCACTCATATTCATTAATGAGCCCATTTTGAAGAAACAATA 300
QY 101 GlnIlyLeuGluThrThrGlyAlaIleValIleLeuAspTyrIleProAspTyrAlaTyrIle 120
Db 301 CAGAAAGCTAGAGACACTCGAGCAAAAGATTCCTGACTCATCCCGATTAAGCATATAT 360
QY 121 ValGluTyrGluGlyAspValGlnSerIysValArgSerIleGluIleValGluSerVal 140
Db 361 GTCAGATATGAGGGGATGTCAATCAAAAGTCCGCTCATTTGAACAGTGAATCAAGTG 420
QY 141 GluProTyrLeuProIleProIleAsnIleAsnProGlnLeuPheThrIysGlyAlaSerThr 160
Db 421 GAGGCAATCTTGGCGAAATACAAATATGATCCCAAGCTTTTCACAAAAAGGCGCATCGAG 480
QY 161 LeuValIlyValAlaLeuAspThrIysGlnAsnAsnValGluValGlnLeuArgGly 180
Db 481 CTGGTGAAGAGCGTTGGCGCTTGATACGAAGCAGAAACATTAAGAAGTGCATTTAAGAGC 540
```

```
QY 181 IleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaIlyPro 200
Db 541 ATTCAGAAATACCGCTCATGTCAGTACGACAAATATCAATGCTTATTAATTAACGCAAGCCT 600
QY 201 GluTyrIysValMetAsnAspValAlaArgGlyIleValIlyValAlaAspValAlaGlnSer 220
Db 601 GAATATTAAGTATGATATGATGATGCGCAGAGATATGTCATTAAGCGGATGTGACAGAGC 660
QY 221 SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
Db 661 AGCTACGGTTTGATATGACAAAGCCAGCATTTGTCAGATGCGCATTCGATGATGATATCA 720
QY 241 GlyArgAsnAspSerSerMetHisGluAlaPheArgGlyIlyIleThrAlaLeuTyrAla 260
Db 721 GGAAGAAACGACAGTTCATGATGCAAGAAAGCTTCCGCGGTAAATAACAGCACTATATGCA 780
QY 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
Db 781 CTGGGTGAGACGAATTAATGCAATGATACGAACGCTCATGTGACCATGTGGCAGGTTCG 840
QY 281 ValLeuGlyAsnGlyAlaThrAsnIlyGlyMetAlaProGlnAlaAsnLeuValPheGln 300
Db 841 GTATTAGGAATGCGCAGCAAGAAATTAAGAAATGCGACCTCAAGCGAATCTGTTTCAA 900
QY 301 SerIleMetAspSerSerGlyIlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe 320
Db 901 TCCATCATGATATGACATGTGGGCTTGAGGCTTGAGGCTTGCTTCCATTCGCAACCTTATTC 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsn 340
Db 961 AGCCAGCATTCATGTCAGTGGCGCAGAAATTCATCAAACTCTGGGGGCGACCGGTGAAT 1020
QY 341 GlyIleTyrThrThrAspSerArgAsnValAspAspTyrValArgIlyAsnAspMetThr 360
Db 1021 GGGGCTACACGACAGATTCAGAAATGTGGATGACTATGAGAAATAATGATATGACG 1080
QY 361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
Db 1081 ATTCCTTTCGCGCTGGGAAATGAAGGCGCAACGCGGTACCAATCAATGCACTCGGTACG 1140
QY 381 AlaAsnAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyr 400
Db 1141 GCTTAAACGCCATTAACAGTGGCGCAACGAAACCTGCGTCCAAAGCTTCGATCTCAT 1200
QY 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgIlyProThrIysAspGlyArg 420
Db 1201 GCAGATATATTAACCAAGTTGCAAGTTCCTTCCCGTGGCCGCAAAAAGATGGCGCA 1260
QY 421 IleIysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAla 440
Db 1261 ATCAAGCCTGATATGCAAGGCGCCAGGACATACATTTTATCAAGCAATCTTCTTTCGA 1320
QY 441 ProAspSerSerPheThrAlaAsnHisAspSerIysTyrAlaTyrMetGlyGlyThrSer 460
Db 1321 CCGATTCCTCTCTTCTGGGGAATCATGACGCAAAATATGCTTATATGGGTGAACCTCC 1380
QY 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValIlyAsn 480
Db 1381 ATGCAACAACCGATTTGGGGGAATGTTGACAGCTCCGATGACATTTTGTGAATAAT 1440
QY 481 ArgGlyIleThrProIlyProSerLeuLeuIlyAlaAlaLeuIleAlaGlyAlaAlaAsp 500
Db 1441 AGAGAAATCACTCTTAAGCCTTCCCTTATGAAGCAGCTTGAATGACAGGTGCTGAT 1500
QY 501 ValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyIlyArgValThrLeuAspIlySer 520
Db 1501 GTTGATTTGGGTTATCCGAACGAAACCAAGATGGGCGGAGTGAACCTCGATTAATTCG 1560
QY 521 LeuAsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnIlyValAlaThrTyr 540
Db 1561 TTGAAGCTTGCATATGTAAGCAATATCAATGACATTCGATCACTGACAAAAGCAGCATAT 1620
```

QY 541 ThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaPro 560
DB 1621 ACCTTAACTGCAACGGCGGCAAGCAATTGAAAATCTCCCTGGTAAGTCGAGTGCCTT 1680
QY 561 AlaSerThrThrAlaSerValThrLeuValAlaAspLeuAspLeuValIleThrAlaPro 580
DB 1681 GCAAGACACTACGCTCTCTGTAACTCCGTGCAATGATTGGATTGGATTACAGACCA 1740
QY 581 AaNGlyThrArgTyValAlaLysAspPheSerAlaProPheAspAspAspAspAsp 600
DB 1741 AACGGAACAGAAATGTGGGAATGACTTCTCAGCACCAATTGACAAATTAAGGAAATGAC 1800
QY 601 ArgAspAspValGluAspValPheIleAspSerProGlnSerGlyThrTyThrIleGlu 620
DB 1801 CCCAATTAAGTAAGAAATGTAATTTAATTGCCCCCAAGTGAACATTAACATTGAG 1860
QY 621 ValGlnAlaTyAspValProValGlyProGlnAspPheSerLeuAlaIleValAsn 639
DB 1861 GTGCAAGCATATTAATGTGCGGTGGACCAAAACTTCTCGTGGCAATTGTGAAC 1917

RESULT 2
AR368116
LOCUS AR368116 1920 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6376227.
ACCESSION AR368116
VERSION AR368116.1 GI:34601777
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1920)
Takaiwa, M., Okuda, M., Saeki, K., Kubota, H., Hitomi, J., Kageyama, Y.,
Shikata, S. and Nomura, M.
Alkaline protease
Patent: US 6376227-A 3 23-APR-2002;
Kao Corporation; Tokyo;
JPK;
FEATURES
SOURCE location/Qualifiers
1.1920
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,14e-169 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x AR368116 (1-1920)

QY 1 MetArgLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrVal 20
DB 1 ATGAGAAAGAAAGAGAGGTGTTTATCTGTTTATCAGCTGCGCAATCTGCGACTGT 60
QY 21 AlaLeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40
DB 61 GCATTAAACAATCCCTCGGCTGGTGAATGCAAGCACTTTTGATCTGGATTTTAAAGAAATT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPhe 60
DB 121 CAACAACAACCGAAGTCAAGGTGTTCTCAAAACAGCAACAACAGTGGCGCTGCATT 180
QY 61 LeuValGluSerGluAspValLysLeuLeuLysGlyLeuLeuLysLeuGluThrVal 80
DB 181 CTGGTGAAGTCTGAAAATGTGAACCTCTTAAGAGATTGCTTAAAGAACTTAAACAGTA 240
QY 81 ProAlaAsnAspLysLeuHisIleValGlnPheAspGlyProIleGluGluThrLys 100
DB 241 CCGCAATTAATTAACCTCAATGTCTCAATTCAATGAGCCCAATTTTGAAGAAACAATA 300

QY 101 GlnLysLeuGluThrThrGlyValAlaLysIleLeuAspTyIleProAspTyAlaTyIle 120
DB 301 CAGAGCTAAGACAACTGAGCAAGAAATTCAGACTACATGCTCGATTATGACATATAT 360
QY 121 ValGluTyGluGlyAspValGlnSerLysValArgSerIleGluHisValGluSerVal 140
DB 361 GTCGAGTAAAGAGGGGAGATTTCAGTCAAAAGTCCGCTCAATTGAAACCTGGAAATCAGT 420
QY 141 GluProTyLeuProLysTyLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
DB 421 GAGCAACTCTCCGAATAACAAATATAGTCCCGAGCTTTTCAAAAAGCCCATGACG 480
QY 161 LeuValLysValAlaLeuAspThrLysGlnAspAspLysGluValGlnLeuArgGly 180
DB 481 CTGGGAAGACGTGGCGCTTGAATACGAAGCAAAACAATAAGATGCAATTAAGAGGC 540
QY 181 IleGluGluIleAlaGlnTyValAlaSerAspAspValHisTyIleThrAlaLysPro 200
DB 541 ATCGAGGAATGCTCAAGTACCTAGCAAGCAATGACGTCAATTATTTACGGCAAGCCT 600
QY 201 GluTyTyLysValMetAspAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
DB 601 GAATATTAAGGTATGATGAATGATGGCCAGAGGTATTGTCAAGCGGATGGCAAGAGC 660
QY 221 SerTyArgLysTyArgLysGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
DB 661 AGCTACGGTTGTATGACAAAGCCAGATGTGCGAGTTGCCGAATCTGGATTGATAC 720
QY 241 GlyArgAspAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyAla 260
DB 721 GGAAGAAACGACAGTTCATGATGATAGACTTCCCGGTAAATTAACAGCATTAATGCA 780
QY 261 LeuGlyArgThrAspAsnAlaAspAspThrAspGlyHisGlyThrHisValAlaGlySer 280
DB 781 CTGGGTCCGACGAATATAGCGAATGATAGCAAGGTCAATGATCCATGTGCAAGTTGC 840
QY 281 ValLeuGlyAspGlyValThrAspLysGlyMetAlaProGlnAlaAspLeuValPheGln 300
DB 841 GTATTAGGAATATGGGCAACGAATTAAGAAATGGAACCTCAAGCAATCTGTTTTTCA 900
QY 301 SerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAspLeuGlnThrLeuPhe 320
DB 901 TCCATCATGATGATGAGATGGTGGCTTGGAGGCTTGCTTCCATCTGCAACCTTATTC 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAspSerTyGlyAlaAlaValAsn 340
DB 961 AGCGAAGCATTCAGTGAAGTGGCCAGAAATTCATAAATCTCGTGGGGGAGCGGTGAT 1020
QY 341 GlyAlaTyThrThrAspSerThrAspValAspAspTyValArgLysAspAspMetThr 360
DB 1021 GGGGCTTACAGACAGATTCAGAAATGTGATGACTATGTAAAGAAATGATATGACG 1080
QY 361 IleLeuPheAlaAlaGlyAspGluAspProAspGlyGlyThrIleSerAlaProGlyThr 380
DB 1081 ATTCTTTTCGGGCTGGGAATGAAGGCCGAACGGGATACCATATGAGCACTGTGACG 1140
QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAspLeuArgProSerPheGlySerTy 400
DB 1141 GCTAAATAACGCAATTAACAGTGGGGCAACGAAAACCTCGGTCAAGCTTGGTTCAT 1200
QY 401 AlaAspAsnIleAspHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArg 420
DB 1201 GCAATTAATTAATTAACAGATGACAGATTCCTTCCCTGGCCCAAAAGATGGGCGA 1260
QY 421 IleLysProAspValMetAlaProGlyThrTyIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAAGCTGATGTCATAGCCCGCAGGACATACATTTTATACGAAAGTCTTCTTGCA 1320
QY 441 ProAspSerSerPheThrAlaAspHisAspSerLysTyAlaTyMetGlyGlyThrSer 460
DB 1321 CCCGATTCCTCCTTCTGGGCGAATCATGACAGAAATATGCTATATAGGTGGAAAGTCC 1380
QY 461 MetAlaThrProIleValAlaGlyAspValAlaGlnLeuArgGluHisPheValLysAsn 480

Db 1381 ATGGCAACACCGGATTTGTCGGGGAATTTGCACGCTCCGTGACCATTTTGTGAAAAT 1440
Qy 481 ArgGlyIleThrProLysProSerLeuLeuYsAlaLeuIleAlaGlyAlaAlaAsp 500
Db 1441 AGAGGAATCACTCTCCCTAGCTTCCCTATTGAAAGCAGCTTGATGACGGTGTGCTGAT 1500
Qy 501 ValGlyLeuGlyYrThrProAsnGlyAsnGlnGlyYrProGlyYrArgValIleLeuAspLysSer 520
Db 1501 GTTGATTTGGGTTCATCCGACGGAACCAAGAGATGGGCGCAGTGAACCTGATTAATACG 1560
Qy 521 LeuAsnValAlaYrValAsnGlyUserSerAlaLeuSerThrSerGlnLysAlaThrYr 540
Db 1561 TTGAACGTTGCTTATGTGAACGAATCCAGTCCCTATCACTAGCCAAAGACCAATAT 1620
Qy 541 ThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValIleTrpSerAspAlaPro 560
Db 1621 ACCTTTACTGCAACGGCGGCAAGCCATTGAATAATCTCCCTGTATGGTCGAGATGCCCT 1680
Qy 561 AlaSerThrThrAlaSerValIleLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
Db 1681 GCAAGCACTACTGCTCTCTGTAACTCGGTCAATGATTTGGATTTGGTCAATTACAGCACCA 1740
Qy 581 AsnGlyThrArgYrValGlyAsnAspPheSerSerAlaProPheAspAsnAsnTrpAspGly 600
Db 1741 AACGGAACAAGATATGTGGGAATGACTTCTCAGCACCATTTGACAAATACCTGGATGGC 1800
Qy 601 ArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyYrThrIleGlu 620
Db 1801 CGCATATAACGTAAAGAAATGATTTATTTATTCGCCCAAGTGAACATATACCATGAG 1860
Qy 621 ValGlnAlaYrAsnValProValGlyProGlnAsnPheSerSerLeuAlaIleValAsn 639
Db 1861 GTGCAAGCAATTAATGTCCGGTGGACCAAAACCTTCGTGTGGCAATTTGTGAAC 1917

RESULT 3
ARS62464
LOCUS ARS62464 1920 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6759228.
ACCESSION ARS62464
VERSION ARS62464.1 GI:53976505
KEYWORDS
SOURCE .
ORGANISM Unknown.
UNCLASIFIED.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6759228-A 3 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;

FEATURES
source Location/Qualifiers
1..1920
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,14e-169 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
DB: 6
US-10-784-870-4 (1-639) x ARS62464 (1-1920)

Qy 1 MetArgLysLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrVal 20
Db 1 ATGGAAGAAAGAAAGAGTGTATTTATCTGTTTATACGTGACGCAATCTGTGCACTGTT 60
Qy 21 AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40

Db 61 GCATTAACAATCCCTGGCGTGGATGCAAGACCTTTGATCTGCAATTTTAAAGAAAT 120
Qy 41 GlnThrThrThrAspValSerGlyPheSerIysGlnArgGlnThrGlyAlaAlaPhe 60
Db 121 CAACCAACAACCGATGTCAAGTGTTCCTCCAAACAGGACCAAAAGAGTGGCGGTGCAATTT 180
Qy 61 LeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLysLysLysVal 80
Db 181 CTGGTGAAGTCTGAAAATGTGAACCTTCTTAAGGATTTGCTAAAGAAACCTTGAACAGTA 240
Qy 81 ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrLys 100
Db 241 CCGGCAATTAATTAATCTCCATATGTCCAAATTCATAGGCCCATTTTGAAGAAACAATA 300
Qy 101 GlnLysLeuGlnThrThrGlyAlaLysIleLeuAspYrIleProAspYrAlaYrIle 120
Db 301 CAGAAGCTAGAGACAACTGAGCAAAAGATTCTCCATACATATCCCTGATTTATGCAATATTT 360
Qy 121 ValGluYrGluGluYrAspValGlnSerLysValArgSerIleGluHisValGluSerVal 140
Db 361 GTCAAGTATGAGGGGGAGATGTTCACTCAAAAGTCCGCTCCATTAACAAGTGAATCAGTG 420
Qy 141 GluProYrLeuProLysYrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
Db 421 GAGGCATACCTTGGCAATACAAATAGATCCCAAGCTTTTCAAAAGGCGCATTCAGCG 480
Qy 161 LeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGly 180
Db 481 CTGGTGAAGAGGTTGGGGCTTGATACAGACGAAACATAAAGATGTCATTTAAGAGGC 540
Qy 181 IleGluGluIleAlaGlnYrValAlaSerAsnAspValHisYrIleThrAlaLysPro 200
Db 541 ATTCAGAAATCGCTCAGTACGTAGCAAGCATACGTCATATATTAACGCAAGCCT 600
Qy 201 GluYrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
Db 601 GAATTAATGAAGATGATGATGTGCGCAGAGTATTTTCAAGCGGATGTGGCACAAGGC 660
Qy 221 SerYrGlyLeuYrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
Db 661 AGCTACGGTTGTATGACAAAGCCCAATGTCTCCAGTGGCCGATTCGATGGAATACA 720
Qy 241 GlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeuYrAla 260
Db 721 GGAAGAAACGACATTCGATGATGAAGCTTCCGCGGTAAATAACGACTATATGCA 780
Qy 261 LeuGlyYrGThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
Db 781 CTGGGTGGAGCAATTAATGCAATGATACAAACGTCATGTGATCCCATGTGGCAGGTTCC 840
Qy 281 ValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
Db 841 GTATTAGGAATGCGCAACGAATTAAGGAATGCGACCTCAAGCGAATCTGTTTTTCAA 900
Qy 301 SerIleMetAspSerSerGlyLysGlyLysLeuGlyLysLeuProSerAsnLeuGlnThrPhe 320
Db 901 TCCATCAATGATACAGATGGGGCTTGAGGGCTTGACCTTGCATTCGCAAACTTATTC 960
Qy 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsn 340
Db 961 AGCCAAACATTCATGTCAGTGCAGGCCAAGATTCATAAATCTCTGGGGGCGAGCGTGAT 1020
Qy 341 GlyAlaYrThrThrAspSerArgAsnValAspAspYrValArgLysAsnAspMetThr 360
Db 1021 GGGGCTTACACGACAGATTCAGAAATGTGGATGATATGTAAAGAAAAGATATGACG 1080
Qy 361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyLysIleThrIleSerAlaProGlyThr 380
Db 1081 ATTTCTTTTGGGGCTGGGAATGAAGGCCGAACGGGTACCAATCAATGCGACTGGGATCG 1140
Qy 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerYr 400

Db	1141	GGTAAAAACGCATTAACAGTCGGCGCAACGAAAACCTCGCTCCAGCTTCGTTCCAT	1200
Qy	401	AlaapaaenllleasnhevalaaglnpbeserarglyprothlyaspolYarg	420
Db	1201	GCAGATTAATATTAACACGTTGGCAGATTCTTCCCGTGGCCGACAAAAGATGGCGGA	1260
Qy	421	IllelybBroapavvalMetAlaProglythrryllleleuseralargSerSerleuAla	440
Db	1261	ATCAAGCCCTGATGTCATGCGCGCAGAGGACATACATTTTATCAGCAAGATCTTCTTCGCA	1320
Qy	441	ProaPserSerPheThrAlaasnhsaPserlytryAlaYrMetGlYglYThSer	460
Db	1321	CCCGATTCTCTCTCTGGGCGCAATATATACACGAAATATGCTTATATGAGTGGAACTCC	1380
Qy	461	MetAlaThrProileValAlaGlYAsnValaaglnleuArgGlwhsPheValYleasn	480
Db	1381	ATGGCAACACCGATTTGTCGGGGAAATGTTCAGACGCTCCGTAGCANTTTGTGAAAAT	1440
Qy	481	ArgGlylllethrProlyPProSerleuLeuYalaaaleuulllealGlyAlaAlaap	500
Db	1441	AGAGAAATCACTCTCCCTTACCTCTCTTATGAAAGACGTTGATTTGCAGGTGCTGCAT	1500
Qy	501	ValGlyleuGlyYrProAsnGlyYAsnGlnGlyTPGlyArgValYThleuAspLySer	520
Db	1501	GTTGGATTGGGTTATCCAGACGGAAACCAAGATGGGGCCAGATGACCTTGATTAATTCG	1560
Qy	521	LeuAsnValAlaYrYValaAsnGlySerSerAlaLeuSerThrSerGlnYalAthrTyx	540
Db	1561	TTGAACGTTGCTATGTACAGAAATCCAGTCCCTCATACATTACGCAAAAAGCAACATAT	1620
Qy	541	ThrPheThrAlaThrAlaGlyLYsPProLeuYlleserleuValTPSerAspAlaPro	560
Db	1621	ACCTTACTGCAACGGCGGGCAGACCATTTGAAAATCTCCCTGATATGGTCCGATGCCCT	1680
Qy	561	AlaSerThrThrAlaSerValYThleuValaAsnAspLeuAspLeuValllethrAlaPro	580
Db	1681	GCAAGCACTACTGCTCTGTAAACCTGTGCATATATTTGGATTTGGTCATTACGACCA	1740
Qy	581	AsnGlyThrArgYrYValGlyAsnAspPheSerAlaProPheAspAsnAsnTPAspGly	600
Db	1741	AAAGGAACAAATATGTTCGGAGATGACCTTCACACACATTTGACATTAACCTGGAGTGC	1800
Qy	601	ArgAsnAsnValGluAsnValPheIlleAsnSerProGlnSerGlyThrYrThrIlleGlu	620
Db	1801	CGCAATTAACGTAGAAAATGTATTTATTAATTCGCCCCCAAGTGGAACATATACATTGAG	1860
Qy	621	ValGlnAlaYrAsnValProValGlyProGlnAsnPheSerleuAlalleValAsn	639
Db	1861	GTGCAGCATATTAATATGTGCGGTTGGACCAAAAACCTTCGTGGCAACATTTGTAAAC	1917

RESULT 4	AR368118	1923 bp	DNA	linear	PAT 12-SEP-2003
LOCUS	AR368118				
DEFINITION	Sequence	7	from patent US 6376227.		
ACCESSION	AR368118				
VERSION	AR368118.1	GI:34601779			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1923)				
AUTHORS	Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y., Shikata,S. and Nomura,M.				
TITLE	Alkaline protease				
JOURNAL	Patent: US 6376227-A 7 23-APR-2002; Kao Corporation; Tokyo, JPY;				
FEATURES					
source	location/Qualifiers				
	1..1923				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					

Alignment Scores:	1,34e-158	Length:	1923
Pred. No.:	3072.00	Matches:	596
Score:	97.0%	Conservative:	23
Percent Similarity:	97.0%	Mismatches:	19
Best Local Similarity:	94.0%	Indels:	0
Query Match:	94.0%	Gaps:	0
DB:	6		

US-10-784-870-4 (1-639) x AR368118 (1-1923)	
QY 2 AAGLYVLYVSLVVAIPIHEUSERVALLEUSERALALALATLEUSERTHRVALA 21	22 LEUAASMPROSERIAGLYASPAIAQTGTHPHEASPLEUASPHELYSGYLIEGLN 41
DB 7 AAGAGAAAAAGAGGTGTTTATCTGTTTATCTGTTTATCTGAGCGAGCTGTTGACCTGGTGG 66	67 TTAAGTAATCCATCTGCAAGTGGTGCAGAGGAATTTGATCTGATCTGATTTCAAGGAATTCAG 122
QY 42 THRTHTHRASPVASERGLYPHESERLYSGINARGINTHRGLYALALALAPHELEU 61	127 ACAACAACAGTAGCTPAAAGGTTTCTCCAAAGAGGGGACAGCTGCTGCTGCTTTCTG 188
DB 62 VALIGUSEGLUASNVALLYSLEULEULYBGLYLEULUALLYBLEUGLUTHRVALPRO 81	187 GTGGAAATCGAAAAATGTGAACCTCCCAAAAGTTTTCAGAGAAAGCTTGAACAGTCCG 246
QY 82 ALAASASNVLYLEUHIISLEVALGINPHEASGLYPROILEUGLUGLUTHRLYSGIN 105	247 GCAAAATATTAATCTCATATTTTTCATTCATTCATGACCAATTTTAGAGAAACAAACAG 306
DB 102 LYBLEUGLUTHRTHGLYALALYSLILEUASPTYRIIPROASPTYRALATYRIIEVAL 122	307 CAGCTGGAAGAAAAACAGGGGCAAGATTCGCACTACATACCTGATTAAGCTTACATTTGTC 366
QY 122 GLUTYRGLUGLYASPVAGLINSERYVALARGSERILEGUHISVALIGUSEVALGLU 141	367 GAGTATGAGGGGAGATTAACTTAAGTCAGCAACAGACCATTTGAGCAGCTGGAGATCCGTGGAG 426
DB 142 PROTYRLEUPROLYSTYRILYSLILEASPPROGLINLEUPHETHRILYSGLYALASERTHLEU 165	427 CCTTATTTGCCATATTAACGAATAGATCCCAAGCTTTTCACAAAGGGGCGATCAGAGCTT 486
QY 162 VALLYVALALEUALALEUASPTHRYLUSGINASASNVLYSGINVALGLINLEUARGYLIE 181	487 GTPAAAGCAGTGGCGCTTGATCAAAAGCAGAAATTAAGAGGTGCATTAACAGGATC 546
DB 182 GLUGLULLEALIEGLINTYVALALASERASNVVALHISTYRIILETHRALALYSPROGLN 201	547 GAACAAATCGCAAAATTCGCATTAAGCATGATGTCCTATATTAATTCGGCAAAAGCCCTGAG 606
QY 202 TYRILYSPVALMETASNPVALALARGGYLIEVALYVALASPVALLAGINSER 221	607 TATAAGGTGATGAATGATGTTGCGGTGAATTTGCAAGCGGATVTGGCTCAGACAGC 666
DB 222 TYRGLYLEUITYRGLINGLYGINLIEVALALAVALLAASPTHRYLLEUASPTHRYGLY 241	667 TACGGGTGTATGAGCAAGAGCAAGATCGTAGGGCTTCCGATACAGGGCTTGATACAGCT 726
QY 242 ARGASNPASPSERSETHIEGLINALAPHEARGLYLYSEILETHRALALEUTYRALALEU 261	727 CGCATATGACAGTTCCATGCATGATGAACCTTCCCGGGGAAATTAATCTCATTAATGATGCAATG 786
DB 262 GLIYAGTTPRANASNVALLASNPSPHYZENGNYHIEGLYTHHISVALALAGLYSERVAL 281	787 GACGCGAAGAAATGACCAATGATACGAATGGTCAATGATGCAATGTCATGCTGCTCCGTA 846
QY 282 LEUGLYAENGLYALATHRASNVLYSEGLYMETALAPROGLINALASNVVALPHEGLINER 301	847 TTAGGAAACGGCTCCATTAATTAAGAAATGGCGCTCAGCGCAATTTAGCTTTCCATCT 906
DB 302 ILEMETASPSERSEGLYGLYLEUGLYLYLEUPROSERASNVLEUGINTHRLEUPHESER 321	907 ATCATGATTAAGCGGTGGGGGACCTTGAGAGCATCACTTCGAATCTGCAACCTTAATTCAGC 966

QY 322 GlnAlaPheSerAlaGlyValArgIleHISerThrAsnSerTrpGlyValAlaValAsnGly 341
Db 967 CAAGCATATCAGTGTGGTGGCGAATTCATACAACTCTGTGGAGACAGATGAT33G 1026
QY 342 AlaTyrThrThrAspSerArgGlnValAspAspTyrValArgValAspAspMetThrIle 361
Db 1027 GCTTACACACAGATTCAGAAATGTGATGACTATGTCCGCAAAATGATATGACGATC 1086
QY 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyIleThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGTCGCGGAAATGAAGACCGAAGCGGAAACCATCATGACACAGGACAGCT 1146
QY 382 LysAsnAlaIleThrValGlyValAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAATGCAATTAACGTGGAGCTACGGAACCTCCGCGCAAGCTTTGGGTCTATGCG 1206
QY 402 AspAsnIleAsnHISValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
Db 1207 GACATATCAACCATGTGGACAGTTCCTTCACTGACCGAACAAGATGAGAGCATC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGATGTCTATGGACCGGGAACGTTCTATCATCATCAGCAAGATCTTCTTGCACCG 1326
QY 442 AspSerSerPheTrpAlaAsnHISAspSerLysTyrAlaTyrMetGlyIleThrSerMet 461
Db 1327 GATTCCTTCCTGGCGCAACCATACAGTAATATGATACATGAGTGAACCGTCCATG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHISpHeValLysAsnArg 481
Db 1387 GCTACACCATGTGCTGTGGAAACGTGGACACAGCTTCGTGAGCATTTTGTAAGAAACGA 1446
QY 482 GlyIleThrProLysProSerSerLeuLeuValAlaLeuIleAlaGlyValAlaAspVal 501
Db 1447 GGCATCAACCAACCAACCTTCTATTTAAAGCGGACATGATGCGGAGCTCAACATC 1506
QY 502 GlyLeuGlyTyrProAsnGlyValAsnGlnGlyTyrGlyValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCCGACCGTAACCAAGATGAGGAGCAATGTGATTAATCCCTCG 1566
QY 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACCTTGCCTATGTGAACAGATCCAGTTCCTCATCCACAGCCAAAAGCGCATCTCG 1626
QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
Db 1627 TTTACTGCTACTGCCGCAAGCCTTTGAAATCTCCCTGTGTGTGATGCCCTGCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACAACCTGCTTCGTGAACGCTTGTCAATGTGAGACTTGTCACTTACCCCTCAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArg 601
Db 1747 GGCACACAGTATGTAGAAATGACTTACTTCCCATACAAAGATACCTGGATGGCCCG 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATAACGTAGAAATATGATTTTATATGACCAACAAGCGGAGCTATCATTAAGAGTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTACCGGTGTGACCAAGAACTTCTCGTTGGCAATGTGAAT 1920
RESULT 5
AR562466
LOCUS AR562466 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 7 from patent US 6759228.
ACCESSION AR562466
VERSION AR562466.1 GI:53976507
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1923)
TAKAIWA,M., OKUDA,M., SAEKI,K., KUBOTA,H., HICOMI,J., KAGEYAMA,Y.,
SHIKATA,S. and NOMURA,M.
TITLE Alkaline protease
JOURNAL Patent: US 6759228-A 7 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;
FEATURES
source Location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-158 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x AR562466 (1-1923)
QY 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAGAAAGAGAGGTTTATCTGTTTATCACTGACGCAATTTGTCACTGTGCG 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTATCATCATGTGACAGTGTGCAAGAAATTTGATCTGATTCGAATTCGAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyValAlaAlaPheLeu 61
Db 127 ACAACAATGATGTCTAAAGGTTTCTCCAGACAGGCGCAGACTGTGTGCTGCTTTCTG 186
QY 62 ValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLysLysLysLysLys 81
Db 187 GTGGATGTGAATATGTAACCTCCCAAAAGATTTTGCAGAAAGATTTGAAACAGTCCCG 246
QY 82 AlaAsnAsnLysLeuHISLeuAlaGlnPheAsnGlyProIleLeuGlnGluThrLysGln 101
Db 247 GCAATATATAAATCCATATTAATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 306
QY 102 LysLeuGluThrThrGlyValAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAGAAACAGGCGCAAGATTCCTCACTACATACCTGATTTATGCTTACATGTGC 366
QY 122 GluTyrGluGluYAspValGlnSerLysValArgSerIleGlnHISValGluSerValGlu 141
Db 367 GAGTATGAGGCGAGTATTAAGTTCAGCAACAACCATTTGACAGTGGAAATCCGTGAG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
Db 427 CCTATTTGGCGGATATACAGATATGATCCCAAGCTTTTCAAAAAGGGGCAATCAGAGCTT 486
QY 162 ValLysValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
Db 487 GTAAAGACAGTGGCGCTTGATACAAAGCAAAATTAAGAGGTGCATTTAAGGGCATC 546
QY 182 GluGluIleAlaGlnIleValAlaSerAsnAspValHISThrIleThrAlaLysProGlu 201
Db 547 GAACAAATCGCAAAATTCGCAATTAAGCATATGCTATATTAATTAACGCAAAAGCTTGA 606
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysValAspValAlaGlnSerSer 221
Db 607 TATAAGGATGAGATGATGATTTGGCGGTGATTTGCAAAAGGAGATGTGGCTCAAGGACGC 666
QY 222 TyrGlyLysTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGTTGTATGCAAGAGCAAGATCCTACCGGTTCGATACAGGCGCTTGATACAGGT 726

QY	242	ArgAsnAAsSerSerMetHeIgeGwIaPheArGgLyVgVtIeThraIaLeuTyraIaLeu	261
Db	727	CGCAATGACAGTTCGATGCAATGAAGCTTCGCGGGGAAATTAATCGCATTAATATGCAATGG	786
QY	262	GIATGThraAsnAsnIaAsnAAsPThraAsnGIyHIGLyThrHiIsvaIaIaGlySerVal	281
Db	787	GGAACGACGAATRAATGCCATGATACATGATGTCATGATGACATATGGCTGGCTCCGTA	846
QY	282	LeuGIyaenGIyAlaThraAsnLySGIyMeAlaAProGInaIaAsnLeuValPheGInSer	301
Db	847	TTAGGAAACGGCTCCCACTAATAAAGAAATGGCGCCTCAGCGCAATCTAGCTTCCAACTC	906
QY	302	IIeMeAAsPSeSerGIyGIyLeuGIyLeuProSerAsnLeuGIInThrIeupheser	321
Db	907	ATCATAGGAAATGGGATGGGGGACCTTGGAGCATCACTTCCAACTCGCAAACTCAATTCAGC	966
QY	322	GIaIaPheSerAlaGIyAlaArGIHeIstThraAsnSerTrpGIyAlaAlaIaIaAsnGIy	341
Db	967	CANCAATACAGTGTGGTGCAGAAATTCATACAACTCTGGGGAGACAGATGAATGGG	1022
QY	342	AlaTyThrThraAsPSeArghaAsnValaAsaPtyrValaArgLyAsnAsnAspMetThrIle	361
Db	1027	GCTTACACAAACAGATTCAGAAATGTGATGATGTCGCAAAATGATATGACGATC	1088
QY	362	LeuPheAlaAlaGIyAsnGIuArProAsnGIyGIyThrIleSerAlaProGIyThraIa	381
Db	1087	CTTTTCGCTCGCGGAAATGAAAGACCGGAAACCATCATAGTCACCAAGCACAGCT	1146
QY	382	LyAsnAlaIaIeThraValGIyAlaThrgIyaMeIuAsArgProSerPheGIySerTyraIa	401
Db	1147	AAAAATGCAATACAGTCCGAGACTACGGAACCTTCGCGCAAGCTTGGGCTTAATGG	1206
QY	402	AspAsnIIeAsnHiIsvaIaIaGInPheSerSerArgLyProThrIyAsnGIyArGIle	421
Db	1207	GACATATCAACCAATGTGGCAAGTCTCTTCACTGACGACCGCAAAAGATGACGATC	1266
QY	422	LySProAspValMeAlaProGIyThrTyrlIeLeuSerAlaArgSerSerLeuAlaPro	441
Db	1267	AAACCGGATGTCATGGCAACCGGAAACGTTCATATCATACAGCAAGATCTTCTTCACCG	1322
QY	442	AspSerSerPheThraIaAsnHiIAspSerLySTyraIaTyMeGIyGIyThrSerMet	461
Db	1327	GATTCCTCTCTGGGCGCAACATACAGTAATATGCAATACATGGGGAACGTCACATG	1386
QY	462	AlaThrProIleValaIaGIyAsnValaIaGInLeuArGIuHIISpHeValLyAsnArg	481
Db	1387	GCTACACCGATGCTGTGGAAACGTGGCAACGCTTCGTAGACATTTTGTGAAAAACGA	1444
QY	482	GIyIleThrProLyAsProSerLeuLeuLybAlaAlaLeuIleAlaGIyAlaAlaAspVal	501
Db	1447	GGCATCACACCAAGCTTCTCTATTAAGGGGACGATGGTCCGGTGCAGCTCAATC	1506
QY	502	GIyLeuGIyTyrrProAsnGIyAsnGIuGIyTrpGIyArGValThrLeuAspLySerLeu	521
Db	1507	GGCCTTGGCTAACCCCAAGCGGTAAACCAAGATGGGACAGATGACATTTGATTAATCCCTG	1566
QY	522	AsnValaIaIaTyrrValaAsnGInSerSerAlaLeuSerThraSerGIyAlaAlaThrTyThr	541
Db	1567	AACGTTCCTATGTAAACAGTCCAGTCTCTATACCAACGCAAAAGCGACGTATCG	1622
QY	542	PheThraIaThraIaGIyLybProLeuLybIleSerLeuValITrpSerAspAlaProIa	561
Db	1627	TTTATCTGCTACGCGGCAAGCTTGTGAAATCTCCCTGTATGTGTGATGCTCCCTGG	1688
QY	562	SerThrThraIaSerValThrLeuValaAsnAspLeuAspLeuValIIeThraIaProAsn	581
Db	1687	AGCAACACTGCTTCGTGAACGTTGTCAATATCGAACCTTGTCAATTAACCCCTCCAAT	1746
QY	582	GIyThraTyrrValaGIyAsnAspPheSerAlaProPheAspAsnAsnTrpAsnGIyArg	601
Db	1747	GGCACACAGTATGTAGAAATATCTTATCTTCGCCATCAATGATTAATCGGGATGGCGCG	1806
QY	602	AsnAsnValGIuAsnValPheIIeAsnSerProGInSerGIyThrTyThrIIeGIyAla	621

```

Db      1807  AATAACGTAAGAAATGATATTATTAATGACACCAAAAGCCGAGCTATTAACATTGAAGTA 1866
Oy      622  Glna1aTyrAnValPProvaIGlyProGlnaSPheSerLeuAlaIleValAsn 639
Db      1867  CAGCTTATAACGTAACCGATTGACACCAAGACTTCTCGTTGGCAATTGGAAAT 1920

RESULT 6
AB051423      1923 bp  DNA      linear  BCT 10-MAY-2002
LOCUS      Bacillus sp. KP43  PRO  gene for protease, complete cds.
DEFINITION  AB051423
ACCESSION  AB051423
VERSION    AB051423.2  GI:20521154
KEYWORDS
SOURCE     Bacillus sp. KSM-KP43
            Bacillus sp. KSM-KP43
ORGANISM   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

REFERENCE   1  Itoh,S. and Saeki,K.
AUTHORS     Itoh,S. and Saeki,K.
TITLE       new protease
JOURNAL     unpublished
REFERENCE   2  (bases 1 to 1923)
AUTHORS     Saeki,K.
TITLE       Direct Submission
JOURNAL     Submitted (21-NOV-2000) Katsuhisa Saeki, KAO, CORPORATION;
            2606, KASABNE, ICHIKAIMACHI, HAGA, TOCHIGI 321-3486, Japan
            (E-mail:3871815@kasanet.kao.co.jp, Tel.8128567471(ex.7471),
            Fax:81285687403)
COMMENT     On May 9, 2002 this sequence version replaced gi:14164344.
FEATURES
             Location/Qualifiers
             1..1923
                /organism="Bacillus sp. KSM-KP43"
                /mol_type="genomic DNA"
                /strain="KP43"
                /db_xref="taxon:109322"
                1..1923
                /gene="PROF"
                1..1923
                /gene="PROF"
                /codon_start=1
                /transl_table=1
                /product="protease"
                /protein_id="BAB5674.2"
                /db_xref="GI:20521155"
                /translation="MRKKKKVPLVSTAAALISTVALSNPSAGANFOLDPFGIOFT
                TDKRSKQOGTCGAAALVSESNVYLPKGLQKLTVPANNKHIIQFNPILEBTKQ
                QLEBTKAKLIDYLPDAIVYIEIGDVKASATSTIEHVSVEPYLPYTRIDQLFKGAS
                ELVAVALDTRQKRKEVQLRGIQIOLAFASNDVLYITAKPEKVNNDVARGIVKDV
                AOSVYGGQGIIVAVADTGLDGRNDSVHGRLFRGKITLVALGRNNANDNGHGT
                HVGASVYGVSTNGMAPQANLTFQSIIMDSGGGLGGPSSLQTLFSDAYSAGARIHTN
                SMGAVNGATTTDSRVNDVDYRKNDMTTLLPAAQNBPNCGTTSAPGTAKNAITVGGATE
                NLRPSRSTADNININVAQSSSRPTQDGRKIPDYMAPGFTILASBSLAPDSSFMANH
                DSKAYGVGTSMAKTPIVAGNVQQLRHFVNNRGIITPEKSLKALAGADIGLQVFN
                GSGWGVNGLDMSKALVAVNBSSTLSQATGSPATAGKPKLISLVSDAPASTTA
                SVTLVNDLDVLTIVAPNGTQYVGNDFTSPYNDMDGNNVENVINAPQSGTYTIRVOA
                YNPVGVQGTSLAIVN"

ORIGIN
Alignment Scores:
Pred. NO.:      2.85e-158      Length:      1923
Score:          3066.00      Matches:      595
Percent Similarity: 96.9%      Conservative: 23
Best Local Similarity: 93.3%      Mismatches: 20
Query Match:    93.8%      Indels:      0
DB:             1      Gaps:      0

US-10-784-870-4 (1-639) x AB051423 (1-1923)
Oy      2  ArglyshYshValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
Db      7  AAGGAAGAAAAGGCGTTTATCTGTTTATVACGTCGACGCGGATTTTGTCGACCTGTGGC 66
Oy      22  LeuLenanPProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLySGIyleGln 41

```

Db 67 TTAAGTATCCATCGAGGTGTCAGAGAAATTGATCGAATTCAAAGAAATTCAG 126
Qy 42 ThrtThrAapValSerGlyPheSerIySglInhArgInThrglyAlaAlaPheLeu 61
Db 127 ACAACACTGATGCTAAAGGTTTCTCCAAAGCAGGGGAGACTGGTGGCTGCTTTCTG 186
Qy 62 ValGluSerGluAenValIyLeuLeuIySglIyLeuLeuIySglIyLeuValPro 81
Db 187 GTGGAATCTGAATAATGTAAACTCCCAAAAGSTTTGCGAAGAACTTGAACAGTCCCG 246
Qy 82 AlaaenAenIySleuHsIleValGlnPheAenGlyProIleleuGluGluThrySglIn 101
Db 247 GCAATATATAACTCATATTATTCAAATTCATAGCAATTTTGAAGAAACAAACAG 306
Qy 102 LysleuGluThrtThrglyAlaIySglIleuAapIyTlleProApyTAlaTyrIleVal 121
Db 307 CAGCTGAAAAAACGGGGCAAAAGATTCGACTACATACCTGATTCATTCATTCG 366
Qy 122 GluTyrGluGluAapValGlnSerIyValArgSerIleGluHsValGluSerValGlu 141
Db 367 GAGATGAGGGGAGATGATTAAGTCAGCAACAGACCATTCAGCAAGTGGAAATCCGTGAG 426
Qy 142 ProTyrLeuProIySglIySglIleAapProGlnLeuPheThrySglIyAlaSerThryLeu 161
Db 427 CCTATTGTCGGAATACAGAAATAGATCCCGAGCTTTACAAAGGGGCAATCAGAGCTT 486
Qy 162 ValIySglIleuAlaIleuAapThrySglIleuAenAenIySglIyValGlnLeuArgIyIle 181
Db 487 GTAAAGAGAGAGGCTTGATTAACAAAGCAAAATTAAGAGGTCAATTAAGAGGCATC 546
Qy 182 GluGluIleAlaGlnIyValAlaSerAenAapValHsIyTyrIleThryAlaIyProGlu 201
Db 547 GAACAATCCGACATTCGCAATTAAGCAATGATGCTATATTTACGGCAAAAGCTTACG 606
Qy 202 TyrIySglIleuAapValAlaIySglIyIleValIySglIyAapValAlaGlnSerSer 221
Db 607 TATTAAGGATGATGATGATGCTGGCGGTGAATTTGCAAGCGAGTGGCTCAAGGACGC 666
Qy 222 TyrGlyLeuTyrGlyGlnIyGlnIleValAlaValAlaAapThrglyLeuAapThryGly 241
Db 667 TACGGGTGTATGACAGAGACAGATCGTACGGTGGCCGATACAGGGCTTGATTAACAGGT 726
Qy 242 ArgAenAapSerSerMetHsIleGluAapThrglySglIyThryAlaIleuTyrAlaIleu 261
Db 727 CGCAATGACATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
Qy 262 GluArgThryAenAenAlaAenAapThrglySglIyThryHsValAlaIleGlySerVal 281
Db 787 GGACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
Qy 282 LeuGlyAenGlyAlaThryAenIySglIyMetAlaProGlnAlaAenIleValPheGlnSer 301
Db 847 TTAAAGAAACGGCTCAATTAAGAAAGAAATGCGCTCAGGGAATTCATGCTTCCAACT 906
Qy 302 IleMetAapSerSerGlyIyLeuGlyIyLeuProSerAenLeuGlnThryPheSer 321
Db 907 ATCATGATTAACGGTGGGGGACTTGGAGGACTACTTGGATTCGCAAACTTATTCAGC 966
Qy 322 GlnAlaPheSerAlaGlyAlaArgIleHsIyThryAenSerTyrGlyAlaAlaValAenGly 341
Db 967 CAACACATCACTGCTGGTGGCAGAAATTCATCAAACTCTGGGGGAGCAGTGAAATGGG 1026
Qy 342 AlaIyThryThryAapSerArgAenValAapApyTyrValAaglyAenAapMetHryIle 361
Db 1027 GCTTACACAAAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATG 1086
Qy 362 LeuPheAlaAlaGlyAenGluArgProAenGlyIyThryIleSerAlaProGlyThryAla 381
Db 1087 CTTTTCCTGCTGGGAAATGAAAGACCAAGCGGCAACATCATGTCACGAGGCAAGCTT 1146
Qy 382 LysAenAlaIleThryValIyAlaThrglyAenLeuArgProSerPheGlySerTyrAla 401

Db 1147 AAAATGCAATTAACATGTCGAGATCAAGAAAACTCCGCCCAAGCTTTGGCTTATTCG 1206
Qy 402 AspAenIleAenHsValAlaGlnPheSerSerArgIyProThryAapGlyArgIle 421
Db 1207 GACAAATTAACAACATGTCGAGATTCCTTCAGGTGACCAAGCAAAAGATGACGAGATC 1266
Qy 422 LysProAapValMetAlaProGlyThryTyrIleuSerAlaArgSerSerIleuAlaPro 441
Db 1267 AACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
Qy 442 AapSerSerPheTyrAlaAenHsIleAapSerIySglIyAlaTyrMetGlyIyThrySerMet 461
Db 1327 GATTCCTCTCTGGGGAACCAATGACATGATTAATCATATCATATGATGATGATGATGATG 1386
Qy 462 AlaThryProIleValAlaGlyAenValAlaGlnLeuArgGluHsPheValIyAenArg 481
Db 1387 GCTACACCGATCGTGTGTCGAAAGTGGACAGCTGTGTGATGATGATGATGATGATGATG 1446
Qy 482 GluIleThryProIySglIySglIySglIySglIySglIySglIySglIySglIySglIy 501
Db 1447 GGCATCACACCAAGCTTCTCTTTAAAGCGGCACTGATTCGGGTGACGTGACATC 1506
Qy 502 GluLeuGlyIyTyrProAenGlyAenGlnIyTyrGlyArgValThryLeuAapIySglIy 521
Db 1507 GGCCTGGCTACCCGAACGGTAAACCAAGATGGGGAAGATGATGATGATGATGATGATGATG 1566
Qy 522 AspValAlaTyrValAenGluSerSerAlaIleuSerThrySglIySglIySglIySglIy 541
Db 1567 AACGTTCCTATGTCGAAACGATCGATCTCTATCCACAGCCCAAAAGGACATCTACG 1626
Qy 542 PheThryAlaThryAlaGlyIySglIySglIySglIySglIySglIySglIySglIySglIy 561
Db 1627 TTTACTCTACTGCGCGGCAAGCTTTGAAATCTCCCTGGATATGATGATGATGATGATGATG 1686
Qy 562 SerThryThryAlaSerValThryLeuValAenAapIySglIySglIySglIySglIySglIy 581
Db 1687 AGCACAACGCTCTCGTACAGCTGTGCAATGATTCGACCTTGCATTAACCGCTCCAAAT 1746
Qy 582 GluThryArgTyrValIySglIyAenAapPheSerAlaProAenAenAenAenAenAenAen 601
Db 1747 GGCACACAGTATGATGAAATGATCTTACTTCGCAATACATGATTAATCTGGGATGGCGC 1806
Qy 602 AspAenValGluAenValPheIleAenSerProGlnSerIyThryTyrIleGluVal 621
Db 1807 AATTAACGTAAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1866
Qy 622 GlnAlaTyrAenValProValGlyProGlnAenPheSerIleuAlaIleValAen 639
Db 1867 CAGGCTTATTAACGTAACGCTTGGACACAGACCTTCTCGTTGGCAATTTGAAAT 1920

RESULT 7
CO888285 1923 bp DNA linear PAT 19-OCT-2004
LOCUS
DEFINITION Sequence 2 from Patent EPI466970.
ACCESSION CO888285
VERSION CO888285.1 GI:54304586
KEYWORDS
SOURCE
ORGANISM Bacillus sp. KSM-KP43
Bacillus sp. KSM-KP43
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Sato, T., Okuda, M., Koyama, S., Iwawa, Y. and Kobayashi, T.
Alkaline protease
Patent: EP 1466970-A 2 13-OCT-2004;
JOURNAL Location/Qualifiers
FEATURES
source
1..1923
/organism="Bacillus sp. KSM-KP43"
/mol_type="unassigned DNA"
/db_xref="taxon:109322"
1..1923
/note="unamed protein product"
/codon_start=1

CDS
1..1923
/note="unamed protein product"
/codon_start=1

```
/crans1 table=11
/protein_id="CAM61612.1"
/db_xref="GI:54304587"
/translation="MKKKKVFSLVLSAAAILSTVALSNPSAGANFDLPKGIQTT
TDAKGFSKQGTGAAPLVESBNVKLPKGIQKLETPVANNKLIHQFQPILEETKQ
OLEKTKAKILIDYIDYAVIVYEGDVKSTSTIEHVSVEPLPIRIRIDPOLLETKAS
ELVKVVALDPTKQKKEVOLRGIEQIAQFISNDVLYTKPKYKWNVDARQIVKADY
AGSYGIVGCGQIVAVADPTGIDPTGRDMSMHPAFKQITRALVALGRTNNANIGHGT
HYAGSVLNGSTKKGMAPOANLVFQSIHQSGGLGGJPSNLQLESQATSAQARLHTN
SWGAAVNGAYTTDSRNVDYVRKNDMTLLPAGNBPNGGTLIAPGTAKNALTVGATE
NLRPSFGSYADNINHVAQPSRPGTDGRIKQDVAPGTFIISARSISLAPDSFVANH
DEKIVAMGGSMTAPIVAGNVAQIREHFVKNQITPKPSILKALIALAGADIGTQPN
GQNGMGRVTLDSKSLNVAYNNESSLSSTQKATYSPATATGKPKLKLISVNSDAPASTA
STTVVNDLDLYITAPNGTQYVGNDFPSFINDWDRNNVNPINAPQSGTTTIEVQA
INVPVQPTFSALIVN"
1.618
619
```

```
Alignment Scores:
Pred. No.: 2,85e-158 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 6 Gaps: 0
```

US-10-784-870-4 (1-639) x CQ888285 (1-1923)

```
QY 2 ARGVLYSLYSLVVALPHEUSERVALLEUSERALALALALEUSERTHRVALA 21
DB 7 AAGAGAAAGAGGTGTTTATCTGTTTATCAGCTGCAGCAATTGTGCAGCTGTGG 66
QY 22 LEUAAANPPOSERALAGLYAPRALAARGTHRPHESRLEUAPHELYSGYLEG 41
DB 67 TTAAGTAATCCATCTGCAGGTGGTGCAAGAAATTTTGATCTGCAATTTCAAAG 126
QY 42 THRTHRTHASRYALSERGLYPHESRLYSGINARGINTHRGLVALAALAPHELE 61
DB 127 ACAACAACGATGCTAAGAGTTTCTCCAAAGCGGCGACAGCTGGTCTGCTTTTC 186
QY 62 VALGUSERGLYASRYALYSLYSEULYSGLYLEULYSLYSLYSLYSLYSLYSLY 81
DB 187 GTGGAACTGAAATGTGAATCTCCAAAGGTTTCAGAAAGAGCTTGAACAGTCCG 246
QY 82 ALAASRYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 101
DB 247 GCAATTAATAATCTCCATATTAATCCAAATGAGCAATTTTGAAGAACAAACG 306
QY 102 LYSLEUGLUTHRTHGLVALYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 121
DB 307 CAGCTGGAAAAACAGGCGCAAGATTTCTGCACTACATACCTGATTAATGCTT 366
QY 122 GLUYTGLUGLYASRYALYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 141
DB 367 GAGTATGAGGGGAGTGTAAAGTCAAGCAACAGCAATGAGGACCTGGAATCCG 426
QY 142 PROTYLLEUPROLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSL 161
DB 427 CCTTATTTGCGGATATACAGATAGATCCCAAGCTTTTCAAAAGGCGCATCAG 486
QY 162 VALLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 181
DB 487 GTAAAAAGCGTGGCGCTGTATCAAGCAAGAAAAATTAAGAGGTGCATTAAG 546
QY 182 GLUGLUILEALGLTYRVALALASERASRYALYSLYSLYSLYSLYSLYSLY 201
DB 547 GAACAATGCAACAAATTCGAATAGCAATGATGTCTATATTAACGCAAGCCTG 606
QY 202 TYTLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 221
DB 607 TTTAAGGTATGATATGATGTGGCGCTGGAATTTGTCAAGCGGATGTGGCT 666
```

```
QY 222 TYRGLYLEUPLYRGLYGLNGLYGLNILEVALAVALAALASPTHRGLYLEUAP 241
DB 667 TACGGGTGTATGACAAAGACAGTACGATGCGGTGCGGATACAGGGCTTGATAC 726
QY 242 ARGAAANPSESESESESETHGLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 261
DB 727 CCGAATGACAGTTCCATGATGATGAGCTTCCCGGGGAAAAATTAATGCAATTA 786
QY 262 GLYATGTHRASRYALASRYALASRYALASRYALASRYALASRYALASRYAL 281
DB 787 GACCGGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 846
QY 282 LEUGLYASRYALASRYALASRYALASRYALASRYALASRYALASRYALASRY 301
DB 847 TTAGGAAACGCTCCCACTAATAAGCAATGCGCTCAGCGCAATCTATCTTCA 906
QY 302 ILEMEASPSERSESGLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSL 321
DB 907 ATCATGATAGCGGTGGGAGACTTGAGAGACTACCTTCCAACTGCAACCTTAT 966
QY 322 GLNALAPHESERALAGLYALARGTLEHETHRASNESTYRGLYALALAVALA 341
DB 967 CAAGCATACAGTGTGGTGCAGAAATTCATACAACTCTGGGAGACAGAGTAAT 1026
QY 342 ALATYRTHRTHASRYALASRYALASRYALASRYALASRYALASRYALASRY 361
DB 1027 GCTTACAAACAGATTCGAAATGTGATGATGATGATGATGATGATGATGAT 1086
QY 362 LEUPHEALALAGLYASRYALARGPROANGLYSLYSLYSLYSLYSLYSLYSLY 381
DB 1087 CTTTTCGCTCCCGGAAATBAGACCGAACCGGAAACCATGATGCACACGACAG 1146
QY 382 LYSAAANALIEHTRVALGLYALATHRGLYASRYALASRYALASRYALASRY 401
DB 1147 AAAAATGCAATTAACAGTGCAGACTACGAAACCTCCGCCCAAGCTTGGT 1206
QY 402 AAPAANILEAHNILEVALALAGINPHESERARGLYPROTHRTHASRYALARG 421
DB 1207 GACAAATACCAATGATGACAGATTCCTTCAACGAGAACCAAGATGACAGAT 1266
QY 422 LYSPROASRYALMEALAPROGLYTHRTHYSLYSLYSLYSLYSLYSLYSLY 441
DB 1267 AAACCGATGATGATGACCGGGAACGTTCAATCATACAGCAAGATCTTCTG 1326
QY 442 AAPSESESESESESESESESESESESESESESESESESESESESESESE 461
DB 1327 GATTCCTCTCTGCGGCAACCATGACAGTAATATGATACATGAGGTGGAAC 1386
QY 462 ALATHRPROILEVALALAGLYASRYALALAGINLEUARGGLYHLEPHEVAL 481
DB 1387 GCTAACCGATGATGATGAGAAACGTGGCAAGCTTCGTAACATTTGTGA 1446
QY 482 GLYILEHTRPROLYPROSELEULYSLYSLYSLYSLYSLYSLYSLYSLYSLY 501
DB 1447 GGCATACACCAAAAGCTTCTTATTAAGCGGCACTGATGCGGCTCAGCTG 1506
QY 502 GLYLEUGLYTYPROANGLYASRYALYSLYSLYSLYSLYSLYSLYSLYSLY 521
DB 1507 GGCCTTGGCTACCGGACGCTAACAGGATGCGGACGATGATGATGATGATG 1566
QY 522 AENVALALATYRVALASRYALASRYALASRYALASRYALASRYALASRYAL 541
DB 1567 AACGTTCCTTAATGTAACAGATGCAAGTCTCATGCAACAGCAAAACGAT 1626
QY 542 PHEHTRALATHRALGLYASRYALASRYALASRYALASRYALASRYALASRY 561
DB 1627 TTTACTGCTACTGCGGCAAGCTTGTAAATCTCCGTGATGATGATGATG 1686
QY 562 SEHTRTHALASRYALASRYALASRYALASRYALASRYALASRYALASRYAL 581
DB 1687 AACAACAGTGTCCGTAACGCTTGTCAATGATGATGATGATGATGATGATG 1746
QY 582 GLYTHRARGTYRVALGLYASRYALASRYALASRYALASRYALASRYALASRY 601
```

|||||.....
Db 1747 GGCACACAGTATGTAGAAATGACTTACTTCCCATACATACATACATCGGATGCCGCC 1806
Qy 602 AaenavajlgluavnaiphelleanserProginSerclyThrTyThrIlegluval 621
Db 1807 AATAACGTAGAAAATGTATTTATTAATGACCAACAGCGGAGCGATATCAATTAGAGTA 1866
Qy 622 GlnalatyrraenvalProvalglyProginanpheneserleualillevalaen 639
Db 1867 CAGGCTTATTAACGTACCGGTGTGACCAACGACCTTCTCGTTGGCAATGTGTGAAT 1920
RESULT 8
CQ889309 1923 bp DNA linear PAT 19-OCT-2004
LOCUS Sequence 2 from Patent EP1466962.
ACCESSION CQ889309
VERSION CQ889309.1 GI:54305176
KEYWORDS
SOURCE Bacillus sp. KSM-KP43
ORGANISM Bacillus sp. KSM-KP43
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Okuda, M., Sato, T., Takimura, Y., Sumitomo, N. and Kobayashi, T.
TITLE Alkaline protease
JOURNAL Patent: JP 1466962-A 2 13-OCT-2004;
Kao Corporation (JP)
FEATURES
Source location/Qualifiers
1..1923
/organism="Bacillus sp. KSM-KP43"
/mol_type="unassigned DNA"
/db_xref="taxon:109322"
1..1923
/note="unassigned protein product"
/codon_start=1
/transl_table=11
/protein_id="CAH6186.1"
/db_xref="GI:54305177"
CDS
translation="MRKKKVFSLVLSAAALISTVALSNPSAGARNPDLPFKIQTT
TDAPKPSKQGTGAAAPLVESENVLPGKLGKLETVANANKLHIQRPGLIBETKQ
QLEKTKALIDTIDYATVYEVSDVKASTIEHVESVEPLPYRIDPQFTKGAS
ELVAVNALDTQKKNVQLRIGIQIAQPAISNDVLYITAKPEYKVMNDVAGIYADV
AOSVGLVGQGIYAVADTGLDGTGNDSSMEAPFGKITTAALYALGRNNADNTHGT
HVASGLNGSTNKMAPOANLVFQISIMSGGLGGLPSNLOTLPSQVSGARLHTN
SMGAVNAGATTTDSRNVDDYVRKNDMTLLPAGNEGPRGGTISARCTAKNLTGATE
NLRFSGYADNININVAOPSSRGPTKDRIDYMAPEFTISASSIAPSSFRAN
DSKAIYMGTSMAITIVAGNVAQLEKHVKNRGITPKPSLEKALIAAGADIGLEPN
GNQGWGRVTLDSLVNAVYNSSLSIQKATYSTATAGKPLKISLVMSDAPASTTA
SVTLVNDLIDLVITAPNGTQYVGNDFTSPYNDMDGRNNVENVFINAPOSCTTIEVQA
NVNPGPQTFSALIVN"
1..618
619..1920
sig_peptide
mat_peptide
/product="unassigned"
ORIGIN
Alignment Scores:
Pred. No.: 2,856-158 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.94 Conservative: 23
Best Local Similarity: 93.34 Mismatches: 20
Query Match: 93.84 Indels: 0
Gaps: 0
US-10-784-870-4 (1-639) x CQ889309 (1-1923)
Qy 2 ArglysllyslvalphelleanserValleuseralaaiaaialeuserThrVala 21
Db 7 AAGAGAAAGAGTGTATTATCTGTTATCACTGCAGCATTTTTCGACTGTGCG 66
Qy 22 LeuananproserAlaglyaspaalargthrrhneapleuaspheyllegln 41
Db 67 TTAAGTAATCCATTCGAGGTGTGCAGGAATTTTGTGATTCGAAAGAAATTCAG 126
Qy 42 ThrThrThrAspValserclYpbeserlysglnaTglnThrGlyAlaalaiahelenu 61

|||||.....
Db 127 ACAACACTGATGCTTAAGGTTTCTCCAGACAGGGGCGAGACTGGTGTGCTTTCTG 186
Qy 62 ValglusergluavnaillyleuleulysglyleuleulyslyleuleulThrValpro 81
Db 187 GTGAAATCTGAATAATGTGAATCTCCCAAAAGGTTTGCAAGAAAGCTTGAACAGTCCCG 246
Qy 82 AlaanenlyleuHillevalGlnphelanenProilleuGlnGlnThrlyleu 101
Db 247 GCAATTAATTAATCTCAATTAATATCCAAATTCATAGGACCAATTTTGAAGAAACAAACAG 306
Qy 102 LysleuGlnThrThrGlyAlaalyeilleuaspTyrlileProaspTyrlileVal 121
Db 307 CAGCTGGAAGAAAACAGGGGCAAAAGATTCTGCATACATCACTGATTAAGCTTATGTC 366
Qy 122 GlnTyrgluglyaspaalGlnserlyvalargserilleGlnHlevalGlnserVal 141
Db 367 GAGTATGAGGCGCATGTTAAGTACAGCAACAGCATTTGAGCAGTGAATCCGTGGAG 426
Qy 142 ProTyrlleuProlyslTyrlylealaspProGlnleuPhenlyleGlyAlaSerThrleu 161
Db 427 CTTATTTGCGGATATACAGAAATGATCCAGCTTTTCAAAAAGGGGCATCAGAGCTT 486
Qy 162 VallysalaleualaleuaspThrlyleGlnaanenlysgluvalGlnleuAarglyle 181
Db 487 GTAAAGACAGTGGCCCTTGATACAAAGCAAAATAAAGGTGCATTTAAGAGGCATC 546
Qy 182 GlnGlnIlealaglnTyrrvalalaseranaspValhistrylleThrAlaLysProGlu 201
Db 547 GAACAAATCGCAATTCGCAATTAAGCAATGATGCTATTAATTAATTCGCAAAAGCTTACG 606
Qy 202 TyrllysalaleuanaspValalaarglylylleVallysalaleuaspValalGlnser 221
Db 607 TATTAAGGATGAATGATGTTGCGGTGAATGTGCMAAGCGAGTGGCTCAGAGCAGC 666
Qy 222 TyrglyleuTyrglyGlnGlyGlnillevalalevalaleaapThrlyleuaspThrly 241
Db 667 TACGGTGTGATGACAAAGACAGATGCTACGGTGTCCGATACAGGGCTTGATACAGCT 726
Qy 242 ArganaspSerSerMetHleGlnalaphaarglylyslleThrAlaLeuTyrrAlaleu 261
Db 727 CGCAATGACAGTTCGATGCATGAGGACCTTCCGCGGAAAATTAATGCTTATTAATGCAAT 786
Qy 262 GlnTyrrThranaanaalaleuaspThrbanGlyhileGlyThrHlevalaleGlySerVal 281
Db 787 GSAGGACGAATTAATGCAATGATACGATGATGATGATGATGATGATGATGATGATGATGAT 846
Qy 282 LeuGlylaenGlyalathrasnlyleGlyMetAlaProGlnalaleuValphellean 301
Db 847 TTAGGAAACGGCTCCACTAATTAAGGAATGCGGCTCAGCGGAATCTTCCAAATCT 906
Qy 302 IleleaspSerSerGlylyleuGlylyleuProseranleuGlnThrleuPheser 321
Db 907 ATCATGATATCGGTGGGGGACTTGGAGACTTCTGAAATTCGAAACCTTATTTCCGC 966
Qy 322 GlnalapheserAlaglyAlaargllethrrhanserTPGlyAlaaleValaenGly 341
Db 967 CAGGATACAGTGTGCGGCAAGATTAATCAAACTCTGGGAGCGAGCTGAATGG 1026
Qy 342 AlatyrrThrThrAspSerArgaspaalaspTyrrValaarglybasnaspMetThrIle 361
Db 1027 GCTTAACCAACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGAT 1086
Qy 362 LeuphealalaglylaenGlnargProanGlylyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGCTGCGGGGAATGAAGGACCGAAACCGGGAACCAATCAATGACGACGACGACGCT 1146
Qy 382 LysanlaliethrvalGlyAlaThrGlnuasnleuArgProserPheGlySerTyrrAla 401
Db 1147 AAAAATGCAATTAACGTGGAGCTTACGAAACCTTCGCGCAAGCTTGGGCTTATTCG 1206
Qy 402 AspaenilleuHillevalalaglnpheserSerarglyProThrlyAspGlyArgIle 421

Db 1207 GACAAATATCAACCATGTGGACAGCTTCTCTTCAGTGGACCGCAAGATGACCGATC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerLeuAlaPro 441
Db 1267 AACCAGGATGTCATGGACCGGAGGATTCATATACAGCAAGATCTCTTGGACCG 1326
QY 442 AspSerSerPheThrAlaAsnHisAspSerLeuTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTTCTGGCGCAACCAATGCAATATGCAATGAGGTGGAGCTCAG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIleAsnArg 481
Db 1387 GCTACACCGATCGTGTCTGGAACGTCGACAGCTTCGTAGCACTTTGTGAAAAACAGA 1446
QY 482 GlyIleThrProIleProSerLeuLeuValAlaIleuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCATCACACCAAGCTTCTCTATTAAGCGCATGATGCGGTGAGTGCAGTCATC 1506
QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGCTACCCGACAGGTAAACAGATGGGACAGTGCATGATTAATCCCTG 1566
QY 522 AsnValAlaTyrValAsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTTCCTATGTGAACAGTCCAGTTCCTATCCACAGCAAAAAGCATGATCTCG 1626
QY 542 PheThrAlaThrAlaGlyIleProLeuLysIleSerLeuValIleThrAspAlaProAla 561
Db 1627 TTTACTGCTACATGCGCGACAGCTTGAATCTCCCTGTATGCTATGCTATGCCCTGCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACATGCTCTCCGTAACGCTTGTCAATGATCTGACCTTGTCACTACCGCTCCAAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnThrAspGlyArg 601
Db 1747 GGCACACAGATAGTGAATATGACTTACTTCCGCAATACATGATTAATCGGATGCGCG 1806
QY 602 AsnAsnValGlnAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATAACGTAGAAAATGATATTATTAATGACCAACAAAGCGGACGTATCAATTGAGCTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATACGTACCGGTGGACCAAGACCTTCTCGTGGCAATGTGAAT 1920

RESULT 9
AR368117 1923 bp DNA linear PART 12-SRP-2003
LOCUS AR368117
DEFINITION Sequence 5 from patent US 6376227.
ACCESSION AR368117
VERSION AR368117.1 GI:34601778
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 5 23-APR-2002;
Kao Corporation; Tokyo;
JPK;

FEATURES
source Location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 2,85e-158 Length: 1923
Pred. No.: 3066.00 Matches: 595
Score: 3066.00
Percent Similarity: 96.9% Conservative: 23

Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x AR368117 (1-1923)

QY 2 ArgLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAGAAAAGAGTGTATTTATCTGTTTATGAGCTGCAGCAATTTGTGACTGTGGC 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTAATCCATCTGCAGGTGTGCAGAGAAATTTGATCGAATTCAGAAAGAAATTCG 126
QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaIlePheLeu 61
Db 127 ACAACATCGATGCTAAAGTTCCTCAAGCGGGGACAGATGCTGCTCTCTTTCTG 186
QY 62 ValGlySerGlnAsnValIleLeuLeuLysGlyLeuLeuValLeuGlnThrValPro 81
Db 187 GTGAATCTGAATAATGTGAACCTCCAAAGGTTTGCAGAAAGCTTGAAACAGTCCCG 246
QY 82 AlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlnThrLysGln 101
Db 247 GCAATTAATTAATCTCATATTTATCCAAATTCATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAACAGGGGCAAGATTCGACTCATATCATATTCATATTCATATTCATATTC 366
QY 122 GlyTyrGlnGlyAspValGlnSerSerValAlaArgSerIleGlnHisValGlnSerValGln 141
Db 367 GAGTATGAGGGGAGATGTTAAGTACAGCAACACACATTTGACAGCTGGAAATCCGTGGAG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyValAspThrLeu 161
Db 427 CCTTATTTGCGATATACAGATATAGTCCCAAGCTTTTCACAAAGGGGCAATCAGAGCTT 486
QY 162 ValIleValAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlnValGlnLeuArgGlyIle 181
Db 487 GTAAAAGAGAGTGGCCCTTGATCAAGCAAGAAAATTAAGAGGTCAATTAAGAGGATC 546
QY 182 GlnGlnIleAlaGlnTyrValAlaAspAsnAspValHisTyrIleThrAlaLysProGln 201
Db 547 GAACAATGCACAATTCGCAATTAAGCAATGATGTGCTATTAATTAAGCGCAAGCCGTGG 606
QY 202 TyrLysValMetAspAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
Db 607 TTTAAGGTGATGAATGATCTTCCGCTGGAAATGTCCAAAGCGGAGTGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTATGACCAAGACAGATCGTAGCGGTGCGATACAGGGCTTATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGlnAlaPheAspGlyLysIleThrAlaLeuTyrAlaLeu 261
Db 727 CGCAATGACAGTTCATCATGATGACCTTCCGCGGAAAATTAATCGCATTAATGATGATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisIleGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGGACGATTAATGCCAATGATCGAATGTGATGTCATGTCAGCTGTGCTGCCGTA 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
Db 847 TTAGAAAACGCGCTCATTAATTAAGAAATGCGGCTCAGCGCAATCTAGTCTTCAATCT 906
QY 302 IleMetAspSerSerGlyGlyLeuGlyIleuProSerAsnLeuGlnThrLeuPheSer 321
Db 907 ATCATGATAGAGGAGTGGGAGCTTGGAGAGCTTCTGAAATCTGCAAACTTATATTCAGC 966
QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsnGly 341
Db 967 CAGCATACAGTGTGCTGCGCAAGATTCATCAAACTCTTGGGAGACAGAGTGAATGGG 1026

Qy 342 AAlaTyrThrThrAspSerThrGluValAspAspTyrValArgLysAsnAspMetThrIle 361
Db 1027 GCTTACACAAACGATTCAGAAATGTGATGTCTGTGGCGGAAAAAGATATGACGATC 1086
Qy 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGCTGCGGAAATGAAAGACCGAACCGACATCATGTCACAGGACAGCT 1146
Qy 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAATGCAATACAGTCGAGCTACGGAAACCTCCGCGCAAGCTTTGGGCTTATGCG 1206
Qy 402 AsnAsnIleAsnIleValAlaGlnPheSerSerAlaGlyProThrLysAsnGlyValGlyIle 421
Db 1207 GACATATCAACCAAGTGGACAGATTCTTTCACCTGGACCCAAAGAGATGGAGGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGATGTCATGGACCGGAAACGTTCACTATACGAAATCTTCTTTCACACCG 1326
Qy 442 AspSerSerPheTyrPalaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTCTGCGGCAACCATGACATGTAATATGCAATGAGTGGAAACGTCACATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlyLysPheValLysAsnArg 481
Db 1387 GCTACACCGATCGTGTGGAAACGTGGACACGCTTCGTGAGCATTTGTGAAAAACGA 1446
Qy 482 GlyIleThrProLysProSerLeuLeuLysAlaIleLeuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCACTACACCAAGCCTTCTCTATTAAGGCGACATGATGGCGGCGACGTACATC 1506
Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCGAAACGTTAACAGATGGGACAGATGACATGGAATTAATCCCTG 1566
Qy 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTGGCTTGTGAACGAGTCCAGTCTCTATCCACCAAGCCAAAGCCGTAACCTG 1626
Qy 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
Db 1627 TTTACTGCTACCTGCGGCAAGCTTTGAATACTCCCTGGTATGGTCTGATCCCTGCG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCACTACTGCTTCGTAACGCTTGTCAATGATCTGACCTTGATTAACGCTCCAAAT 1746
Qy 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnThrAspGlyArg 601
Db 1747 GGCACACAGTATGTAAGAAATGACTTCTTCGTCATCAATGAATACCTGGATGGCGCG 1806
Qy 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATTAACGTAAGAAATGATTTTATTAATGACACCAAGCGGACGTAATTAAGATTA 1866
Qy 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTACCGGTGGACCAACGACCTTCTCGTTGGCAATGTGAT 1920

RESULT 10
ARS62465 LOCUS ARS62465 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 5 from patent US 6759228.
ACCESSION ARS62465
VERSION ARS62465.1 GI:53976506
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,

TITLE Shikara,S. and Nomura,M.
JOURNAL Alkaline protease
Patent: US 6759228-A 5 06-JUL-2004;
Kao Corporation; Tokyo;
JFX;
FEATURES location/Qualifiers
source 1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,85e-158 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 6 Gaps: 0
US-10-784-870-4 (1-639) x ARS62465 (1-1923)
Qy 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAAAGAAAAGGTGTTTATCTGTTTATCTGTTTATCTGACCTGACGAGATTGTGCACTGTTGCG 66
Qy 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTATCCATCTGCGAGGTGGTGCAGAAATTTGATCTGGAATTTCAAGAAATTCAG 126
Qy 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPheLeu 61
Db 127 ACAACACTGATGCTAAAGTTTCTCCAGCAGGCGGCACTGATGCTGCTTTCTG 186
Qy 62 ValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLysLysLysLysLys 81
Db 187 GTGGAATCTGAAATGTGAATCTCCCAAAAGTTTGCAGAAAGCTTGAAGACATGCCG 246
Qy 82 AlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlyThrGln 101
Db 247 GCAATATATTAACCTCCATTTATTCATATCAATGACCAATTTTGAAGAAACAAACAG 306
Qy 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAAAACAAGGCGCAAAATCTGACATACATCACTGATTAATGCTTATGCTC 366
Qy 122 GluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGluSerValGlu 141
Db 367 GAGTATGAGGCGATGTTAATGTCAGCAACCAAGCACATTTGACGCGATCCGTGAG 426
Qy 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
Db 427 CTTATTTTCCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGCGATCAGAGCTT 486
Qy 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
Db 487 GTAAGAGAGTGGCGCTTGAATCAAGACAGAAATTAAGAGTGCATTAAGAGGATC 546
Qy 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysProGlu 201
Db 547 GAAACAAATCCGACATTCGCAATTAAGCATGATGTGATATTAATTAACGCAAAACCTGAG 606
Qy 202 TyrLysValMetAsnAspValAlaArgGlyLysValLysAlaAspValAlaGlnSerSer 221
Db 607 TATTAAGGATGATGATGATGTTGCGCGTGAATTTTCAAAGCGGATGTGCTCAAGAGCAGC 666
Qy 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGTTTATGACAAAGACAGATCTAGCGGTGCGGATACAGGGCTTGATACAGCT 726
Qy 242 ArgAsnAspSerSerMetHisGluAlaPheArgLysLysIleThrAlaLeuTyrAlaLeu 261
Db 727 CGCAATGACAGTTCGATGATGAAGCTTCCGCGGAAAAATTAATGCAATTAATGCAATTG 786

QY	262	GIYAQTTHAsbnaa1aasnaaPThraengiyhaieglYThrh1aValaaglyserVal	281
Db	787	GGACGGACGAATATATCCCAATATATCAAGATGGTCAATGACGATGTGGCTCCGTA	846
QY	282	LeuGIYAaSG1YAlaThraenlySG1MeTAlaProGIAlaAenueValPheGInser	301
Db	847	TTAGGAAAGCGCTCCACTAATTAAGAAATGGCCCTCAAGCGAATTAAGTCTTCAATCT	906
QY	302	11emeAasPserSerGIyleuGIyleuProSerAenueGInThraePheSer	321
Db	907	ATCATGATAGCGCGTGGGGGACTTGGAGACATCTTCGAATCTGCAACCTTAATTCAGC	966
QY	322	GIAlaAphSerAlaGIYAlaarg1Ieh1eThraenSerTTpGIYAlaAlaValaengly	341
Db	967	CAAGCATATAGTGGTGTCCAGAAATTCATACAAATCTCGGGAGACAGATGATGG	1020
QY	342	AlaTYThThraApsSerArgaenValaAapTYrValArgLYsaAasPheThrIle	361
Db	1027	GCTTACACAACAGATTCCAGAAATGGATGACTATGTCCGAAATATATTAACATC	1088
QY	362	LeupheAlaAlaGIYAaenGluArgProaenGIYThrIleSerAlaProGIYThrala	381
Db	1087	CTTTTGTGCGCGGAATGAAGACCGAACATCAGTGCACAGGACAGCT	1144
QY	382	LYsaAaAlaIleThrValGIYAlaThrGluAenLeuArgProSerPheGIYserTYala	401
Db	1147	AAAAATGCAATACAGTCCGAACTACGGAAAACTCCCGCCAAAGCTTGGGTCTTAATGCG	1206
QY	402	AaPaenIleAaenIeValaIaGIaPheSerSerArgG1ProThThraApsGIYargIle	421
Db	1207	GACAAATATCAACATGTGGCAAGTCTTTCACGTGACCAACAAAGATGACGAGTC	1266
QY	422	LYsProaPValMeTAlaProGIYThrYIleLeuSerAlaArgSerSerLeuAlaPro	441
Db	1267	AAACCGGATGATGCGACACGGGAAGCTCATATCATACGAAGATCTTCTTCCACCG	1322
QY	442	AaPserSerPheThralaAaenIleAaPserLYrYAlaTYrMeGIYGIYThrSerMeC	461
Db	1327	GATTCCTCTCTTGGCGCAACCATACAGTAATATGATACATGCGGTGGAACTGCCATG	1386
QY	462	AlaThrProIleValaIaGIYAaenValaIaGIaLeuArgGluA1PheValLYsaAarg	481
Db	1387	GCTACACCGATGTGGTGGAAACGTGGCACAGCTTCGTACATTTGTGAATAACAGA	1446
QY	482	GIYIleThrProLYsProSerLeuLeuYsaAlaLeuIleAlaGIYAlaAlaAapYAl	501
Db	1447	GGCATCACACCAAGCTTCTCTATTAAGAAGCGCATGATTGGCGGTGCAGCTCACATC	1506
QY	502	GIYleuGIYTYrProaenGIYAaenGluYTPGIYAlaGValThThraApsLYsSerLeu	521
Db	1507	GGCTTGGCTAACCCCAAGCGTAAACCAAGATGGGACAGATGACATTTGAATCCCTCG	1566
QY	522	AaenValaIaTYrValaAaenGluSerSerAlaLeuSerThraSerGInLYaAlaThrTYrThr	541
Db	1567	AACGTGGCTATGTAACGATGCCAGTCTCTATCACACAGCAAAAAAGCAGATGATCG	1626
QY	542	PheThralaThraAlaGIYLYsProLeuYsaIleSerLeuValaITPpSerAAlaProAla	561
Db	1627	TTTACTGCTACGCGGGAGGCTTTGAAATCTCCCTGGTAATGTCTCGATGCCCTTCG	1686
QY	562	SeThrThralaAserValThThraeValaAaenApsLeuValIleThThralaProAaen	581
Db	1687	AGCAACATCGCTTCGTAACGTTGTCAATGATCTTGACCTTGTCAATACCGCTCCAAAT	1746
QY	582	GIYThraArgTYrValaGIYAaAaPheSerAlaProPheAaPaenAantTpaapGIYArg	601
Db	1747	GGACACAGATGTAAGAAATGACTTATCTTCGCATACAAATGAATACCTGGAGATGGCCG	1806
QY	602	AaenAaenValaGluAaenValPheIleAaenSerProGInserGIYThrTYrThrIleGIYAla	621
Db	1807	AATAACGTATGAATAATGTAATTTATTAATGACACCAAAAGCGGACCTATATCAATTGAGTA	1866
QY	622	GIAlaTYrAaenValProValaGIYProGluAaenPheSerLeuAlaIleValaAaen	639

```

Db          1867 CAGGCTTAAACGTACCGCTTGACACGACGACCTTCCTGCTTGCGCAATGTGAAT 1920
RESULT 11
LOCUS       AB084155                      1923 bp      DNA          linear      BCT 02-SEP-2003
DEFINITION  Bacillus sp. KSM-9865 gene for protease, complete cde.
ACCESSION   AB084155
VERSION      AB084155.1  GI:34392386
KEYWORDS
SOURCE      BACILLUS sp. KSM-9865
ORGANISM    Bacillus sp. KSM-9865
REFERENCE   1  Bacillus sp. KSM-9865
AUTHORS     Okuda,M., Saeki,K. and Kobayashi,T.
TITLE       Bacillus sp. KSM-9865 protease gene
JOURNAL     Published Only in Database (2003)
REFERENCE   2  (bases 1 to 1923)
AUTHORS     Okuda,M., Saeki,K. and Kobayashi,T.
TITLE       Direct Submission
JOURNAL     Submitted (18-APR-2002) Mitsuyoshi Okuda, Kao corporation,
              Biological Science Laboratory, Ichikimachi Akabane 2606, Haga,
              Tochigi 321-3497, Japan (E-mail:okuda.mitsuyoshi@kao.co.jp,
              Tel:81-285-68-7543, Fax:81-285-68-7547)
              Location/Qualifiers
FEATURES             1..1923
                     /organism="Bacillus sp. KSM-9865"
                     /mol_type="genomic DNA"
                     /strain="KSM-9865"
                     /db_xref="taxon:192495"
                     1..1923
                     /codon_start=1
                     /transl_table=1
                     /product="protease"
                     /protein_id="BAC82522.1"
                     /db_xref="GI:34392387"
                     /translation="MRKKKKVFLSVLSAAAILSTVALSNPSAGGARNPDDPFKGIQIT
                     TDAGKFSKQCGTGAALFLVSEENVKLPKGLKLETPVANKLHIVQNPPILEETKQ
                     QLEBKAKILIDYIPDAIVIVSEYEDVSAISTEDVSVPEPIYVIDPOLFKGAS
                     ELVAVATLDMQKKVEQLRGIEQLPAPASINDVLTITADDEKVMNDVARGIKVDV
                     AQSSEGLXGQGIYAVADDTGLDTRNDSSMHEAFRGKITLALYALGRNNANDTRGECT
                     HVAQSVLNGNSTNKGMAPQANLVQSIIMDSGGIGSLPSNLQTLPSQASAGARLITN
                     SWGAVNGAYITDTSRNVDDVVRKNDMTILPAAGNEGNGGTLISAPGTAKARLIVGATE
                     NLARSFSEYADNINHVAFQSPFSRPGTKGRILKPDVMAQFTILARSLSADSFMAH
                     DSKAVKGTGSMATPIVAGVVALREHFVKNRGITTPKPSLLKALILAGADAGGYNN
                     GNOCKGRVTLIDKSLNVAVYVNESSLTSQKATYSPTATACKPLKILVSGDAPGSTTA
                     SVTLVNDLIDVITTPKGTQYVGNDFTSPTANNWDGRNNVENFTINAPQSGTITTEVOA
                     INVEVGPQTSLSLVN"
ORIGIN
Alignment Scores:
Pred. No.:      1.46e-157                      Length:      1923
Score:          3053.00                      Matches:      594
Percent Similarity: 96.4%                  Conservative: 21
Best Local Similarity: 93.1%                Mismatches:  23
Query Match:    93.4%                      Indels:       0
DB:             1                          Gaps:         0
US-10-784-870-4 (1-639) x AB084155 (1-1923)

QY      2  ArglylsylysluysValPheleuserValIeusSerAlaIaIaleIeusErThrrValAla 21
DB      7  AAGAAAGAAAAGCGTCTTTTATCTGCTTTATTCAGCTGCACACGATTTTGTGACATGTTGCG 66
QY      22  LeuanaAnProSerAlaGlyAapAlaAyGThrPheAspLeuAspPheIyGlyIleGln 41
DB      67  TTAAGTATCATCTATCGACGTGCGCAAGAAATTTTGATCTGCAATTTCAAGAGAAATTCAG 126
QY      42  ThrThrThraPValserGlyPheSerIyGlyAargIlnrGlyAlaIaIaIaPheIu 61
DB      127  ACAATTAACTAACTGTTAAAGGTTTCTCCACACAGAGGGGAGAGCTGATGCTGTCGATTTTCG 186
QY      62  ValGluSerGluAnValIyLeuLeuIyGlyLeuLeuIyGlyLeuGluIthrValPro 81

```

```
Db 187 GTGGATATGAAAATGTAATCTCCCAAAAGTTTGCAAGAAACCTTGAAACATCTCCCG 246
Qy 82 ALaasnaenlyLeuHieIleValGlnpheanGlyProIleLeuGlnGlnThrySglN 101
Db 247 GCAAAATATTAACCTCCATATTGTCCCAATTCATATGACCAATTTTGAAGAAACAAACAG 306
Qy 102 LysLeuGlnThrThGlyValIleValIleLeuSerPyrIleProArgPyrIleValIleVal 121
Db 307 CAGCTAGAAAACAGGGGCAAAAGATTCCTCAGCTACATACCTGATTTATGCTTACATCTTC 366
Qy 122 GlnTyrgLnGlyAspValGlnSerLysValArgSerIleGlnHieValGlnSerValGln 141
Db 367 GAGATGAGGGGAGATTTAAGTCACACCAACAGCACCATTGAGAGAGCTGGAAATCCCTAAG 426
Qy 142 ProTyLeuProLysTyIleValIleAspProGlnLeuPheThryLeuGlyValAserThry 161
Db 427 CCTTATTGCTCCGATATACAGATATAGTCCCGAGCTTTTCACAAAAGGGGCAATCAGAGCTT 486
Qy 162 ValIyValIleValIleLeuSerThrySglNAsnaenlyGlnValGlnLeuArgGlyIle 181
Db 487 GTAAAAGCAGTGGCGCTGATACAAATCAAAAATTAAGAGGTCCAAATTAAGAGCAGTC 546
Qy 182 GlnGlnIleValGlnTyIleValIleAserAsnaenValHieTyIleThryIleThryProGln 201
Db 547 GAGCAAAATCGCAATTCGCAACCAAGCAATGATGTCTAATATTAACGGCAAAAGCTGAG 606
Qy 202 TyIyValMetAsnaenValIleArgGlyIleValIyValIleAspValIleGlnSerSer 221
Db 607 TATAATGATGATGAATGATGTTCAGCTGAAATTTGCAAAAGCGAGTGGCGAGAGCAGC 666
Qy 222 TyrgIyLeuTyrgIyGlnGlyGlnIleValIleValIleAspThryIleLeuSerThryGly 241
Db 667 TACGGGTGTGATGAGCAAGAGACAGATCGTAGCGGTGCGGATACAGGGCTTGATACAGGT 726
Qy 242 ArgAsnaenSerSerMetHieGlnIleAspAspThryIleThryIleThryIleValIle 261
Db 727 CGCAATGACAGTTCATGATGATGACCTTCGGGGGAAAATTAATGCAATTAATGATGATG 786
Qy 262 GlnTyrgThryAsnaenValIleAspThryIleValIleThryHieValIleValIleVal 281
Db 787 GAGCGAGAGAAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
Qy 282 LeuGlyAsnGlyValIleThryAsnlyGlyIleMetIleProGlnIleAsnLeuValIleGln 301
Db 847 TTAGCAAAAGGCTCCATTAATTAAGAAATGGCGCTCAGGCGCAATCTAGCTCTCAATCT 906
Qy 302 IleMetAspSerSerGlyIyLeuGlyIyLeuProSerAsnLeuGlnThryLeuPheSer 321
Db 907 ATCATGATAGCGCGTGGGGGACTTGGAGGACTACCTTGGAATCTGCAAAACCTTATTCAGC 966
Qy 322 GlnIleAspSerIleGlyValIleArgIleHieThryAsnSerThryValIleValIleVal 341
Db 967 CAAGATATCAGTGTGCTGGGCGCAAGATTCATCAAACTCTGGGGGAGCAGTGGATGGG 1026
Qy 342 AlaTyThryThryAspSerArgAsnaenValIleAspPyrIleValIleValIleValIleVal 361
Db 1027 GCTTACACACAGATTCAGAAATGTGATGCTATGTGCGGAAAATGAATATGCGATC 1086
Qy 362 LeuPheValIleValIleValIleValIleValIleValIleValIleValIleValIle 381
Db 1087 CTTTTCGCTGCGGGAATGAAGAGACGAAACGGCGGAACATGAGTGCACAGGCGCAGCT 1146
Qy 382 LysAsnaenValIleThryValIleValIleThryIleValIleValIleValIleValIle 401
Db 1147 AAAAATGCAATTAAGTGTGAGCTTCAGGAAACCTCCCGCCAAAGCTTGGGCTTATCG 1206
Qy 402 AspAspIleAsnHieValIleValIleValIleValIleValIleValIleValIleVal 421
Db 1207 GACATATACACATGATGAGCAGATGCTTCTTCACTGAGACGACAAAGAGATGAGAGATC 1266
Qy 422 LysProAspValMetIleProGlyThryIleLeuSerIleAspSerSerLeuIlePro 441
```

```
Db 1267 AAACCGATGTCATGAGCAGCGGAAACGTTTCATATCAGCAAGATCTTCTTGACCG 1326
Qy 442 AspSerSerPheThryIleValIleAspSerLysTyIleThryMetGlyIyThrySerMet 461
Db 1327 GATTCCTCTCTTGGGGGAACATGACAGTAAATATCATATCAGGTGGAGACCTCCATG 1386
Qy 462 AlaThryProIleValIleValIleValIleValIleValIleValIleValIleValIle 481
Db 1387 GCTACACCGATGCTTCTGGAACCTGGACACGCTTGTGAGCATTTTGTGAAAAACGA 1446
Qy 482 GlyIleThryProLysProSerLeuLeuValIleValIleValIleValIleValIleVal 501
Db 1447 GGCATCACACCAAGCCCTTCTATTAAGCGGCACTGATGCTCGGGGAGCAGACATC 1506
Qy 502 GlnLeuGlyTyIleProAsnGlyAsnGlnGlyTyIleValIyValIleThryLeuAspLys 521
Db 1507 GGCCTTGCTTACCGCAACCGTAAACGAGATGGGGAGAGATGAGATGGATTAATCCCTA 1566
Qy 522 AsnValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 541
Db 1567 AACCTTGCTTATGTAACGAGTCCAGATCTTATTCACACGCAAAAGCAATGATCTCG 1626
Qy 542 PheThryIleThryIleValIyProLeuValIleSerLeuValIleThrySerAspValIle 561
Db 1627 TTACTGCTACTGCGGCAAGCCTTTGAAGATCTCCCTGATGATGCTGATGCCCCCTGCG 1686
Qy 562 SerThryThryIleSerValIleThryValIleAsnaenLeuValIleIleThryIlePro 581
Db 1687 AGCAACAATGCTTCCGTAACCGCTTGTCAATGACCTTGACCTTGCATTAACGCTCAAT 1746
Qy 582 GlnTyrgThryValIleValIleValIleValIleValIleValIleValIleValIle 601
Db 1747 GGCACAAATATGTTGAAATGATGATTAATTCCTGCGCAATCAATTAATGATGATG 1806
Qy 602 AsnaenValIleValIleValIleValIleValIleValIleValIleValIleValIle 621
Db 1807 AATAAGCTAGAAATGATTAATTAATGCGCAAAAGCGGAGATTAATGATGATGATG 1866
Qy 622 GlnIleThryAsnaenValIleProValIyProGlnIleAsnPheSerLeuValIleVal 639
Db 1867 CAGGCTTATTAACGTAACCGGTTGACCAACAGACCTTCTCGTTGGCAATTTGTAAC 1920

RESULT 12
LOCUS AR069954 3003 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 41 from patent US 5891701.
ACCESSION AR069954
VERSION AR069954.1 GI:7220842
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
1. .3003
/organism="unknown"
/mol_type="unassigned DNA"

ALIGNMENT Scores:
Pred. No.: 1.75e-143 Length: 3003
Score: 2799.00 Matches: 541
Percent Similarity: 92.4% Conservative: 52
Best Local Similarity: 84.3% Mismatches: 45
Query Match: 85.7% Indels: 4
DB: 6 Gaps: 2

US-10-784-870-4 (1-639) x AR069954 (1-3003)
Qy 1 MetArgLys-----LysIyValIlePheLeuSerValIleuSerValIleValIleValIle 17
```

```

Db      846 ATGAGAAAGAAAGATCGAAGGGTTTTTTATCGCTTTATCATGTTGTCGACATATG 905
Qy      18 SerThValAlaLeuAenPProSerAlaGlyAaspAlaArgThrPheApleAaspPhe 37
Db      906 TCTTGTGTTGCTTTAAGCAGCTCTTCTATATGCGGCGAAACAATTGGAATTGGACTTT 965
Qy      38 LysGlyIleGlnThrThrThrAaspValSerGlyPheSerLysGlnArgGlnInThrGlyAla 57
Db      966 AAGGGATGAGACACTTACGCTAGAGAGCGGCCCAAGCAAGCAAGAAAGCGGAGAG 1025
Qy      58 AlaAlaPheLeuValGluSerGluAenValLysLeuLysGlyLeuLysLysLeu 77
Db      1026 GCATCTTTCTTTGAAACTCTGAAAATGTGAAAATCCCAAGAGATTCACAAAAGAACTA 1085
Qy      78 GluThrValProAlaAenAenLysLeuHisIleValGlnPheAengLysProIleLeuGlu 97
Db      1086 GAAGTGGTTCACGCGATTAACAAGCTATATATGCTTCAATTGACGGAACCTATTTAGAG 1145
Qy      98 GluThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAaspTyrIleProAaspTyr 117
Db      1146 GAAGCCAACTTCACTACAGAAAGCGGAGCGAAATTCGATTAACAACAGATTAAC 1205
Qy      118 AlaTyrIleValGluTyrGluGluYAspValGlnSerLysValArgSerIleGluHisVal 137
Db      1206 GCTTATATTTGCGAATATGATGCGGATGTAAAGCCGTAACCTAACCGCAATTTGCCATTGG 1265
Qy      138 GluSerValGluProTyrLeuProLysTyrLysIleAaspProGlnLeuPheThrLysGly 157
Db      1266 GAATCGGTGAAACCAATTTACCTTATATTAATAGACCCGCAATTAATTTCCAGAGGA 1325
Qy      158 AlaSerThrLeuValLysAlaLeuAlaLeuAaspThrLysGlnAenAenLysGluValGln 177
Db      1326 GCTTCAAAATAGTAAACAGTAGCTTTAGATTAAGAAAGCAAGAAAGAAAGTAACGT 1385
Qy      178 LeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAenAenValHisTyrIleThr 197
Db      1386 TTAAGAGATTGGAACAATTTGCCCAATACCGCAACAATATATATATATATATATATATAT 1445
Qy      198 AlaLysProGluTyrLysValMetAenAenValAlaArgGlyIleValLysAlaAaspVal 217
Db      1446 CCAAGGCCGAAATCAAGATTTTGAATGACGTGCGCCGATGCGCATTTGAAAGCAGACGTC 1505
Qy      218 AlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAaspThrGly 237
Db      1506 GCACAAAATTAACCTTTGCTTATATGGAACAAGACAATGTGTAGCAGATTGCTGATCTGGG 1565
Qy      238 LeuAaspThrGlyAArgAenAenSerSerMetHisGluAlaPheArgGlyLysIleThrAla 257
Db      1566 CTTGATACAGAAAGAAATGACAGTTGATGATGAATTCGCGCGATGAATTAACCGCA 1625
Qy      258 LeuTyrAlaLeuGlyAArgThrAenAenAlaAenAaspThrAenGlyHisIleThrHisVal 277
Db      1626 CTATATGCACTGGGCGAAGCAATACGCCAATGATCCAATGGAATGGAATGGAATCCCATGTT 1685
Qy      278 AlaGlySerValLeuGlyAenGlyAlaThrAenLysGlyMetAlaProGlnAlaAenLeu 297
Db      1686 GCTGATCTGTGTGGAAT---GCTACAAATTAAGGATGCGACCGCAAGCAATCTTA 1742
Qy      298 ValPheGlnSerIleMetAenAenSerSerGlyLysGlyLysIleLeuProSerAenLeuGln 317
Db      1743 GCTTTCAATCTATATATGATATGATGAGAGCGCTGGAGAGACTCTCTATATCTACAA 1802
Qy      318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisIleThrAenSerTyrGlyAla 337
Db      1803 ACATTTCTCAGTCAAGCATATAGTCTGGAGCGAAGATTCATACGAATTCATGGGGGGCT 1862
Qy      338 AlaValAenGlyAlaTyrThrThrAaspSerAArgAenValAaspTyrValArgLysAen 357
Db      1863 CCAAGTAACGGTGGCTATACGACGACTTCGAAAGTGAATATATATATATATATATATATAT 1922
Qy      358 AaspMetThrIleLeuPheAlaAlaGlyAenGluAArgProAenGlyLysIleThrIleSerAla 377

```

```

Db      1923 GATATGACGATTTCTTTTGGCGCGGAATGAGGGAACAAGTAGCGGTACATCACTGCA 1982
Qy      378 ProGlyThrAlaLysAenAlaIleThrValGlyAlaThrGluAenLeuArgProSerPhe 397
Db      1983 CCAAGAAACAGCAAAAATGCGATTACAGTTGGGCGAACCGAAACCTACGTCACAGCTTC 2042
Qy      398 GlySerThrAlaAenAenIleAenHisValAlaGlnPheSerSerTyrGlyProThrLys 417
Db      2043 GAGTCTTAATGCGATATATATTAACAGATGTTGCTCAATTTCTTCAAGAGGTCTTACTAGA 2102
Qy      418 AaspGlyArgIleLysProAaspValMetAlaProGlyThrTyrIleLeuSerAlaArgSer 437
Db      2103 GATGACAGATTTAAGCGGACGTCATGAGCACAGAGTACGATATTTCTCTGCTATATCA 2162
Qy      438 SerLeuAlaProAaspSerSerPheTyrAlaAenHisAaspSerLysTyrAlaTyrMetGly 457
Db      2163 TCATTAGCTCCAGATTCCTCATTTCTGGGCAACCAAGATATGTAATATGCTCAATGGGT 2222
Qy      458 GlyThrSerMetAlaThrProIleValAlaGlyAenValAlaGlnLeuAargGluLysPhe 477
Db      2223 GGTACTTCTATGCTACTCTCAATTTAGAGAGTAAATGTTGCACAAATTAAGGAGCATTTT 2282
Qy      478 ValLysAenAargGlyIleThrProLysProSerLeuLeuLysAlaLeuIleAlaGly 497
Db      2283 GTGAAAATAATAGAGGCTAATCTCTAAGCTTCCCTTTAAAGCTCTTAAATTCAGAGT 2342
Qy      498 AlaAlaAaspValGlyLeuGlyTyrProAenGlyAenGlnGlyTyrGlyValThrLeu 517
Db      2343 GCTGCGAGATGTTGACTTGGCTTTCCAAATGTGTAAACAAGATGGGGAAGATGACGTTA 2402
Qy      518 AaspLysSerLeuAenValAlaTyrValAenGluSerSerAlaLeuSerThrSerGlnLys 537
Db      2403 GATTAATCCCTAATATGTCGATTTGTGAATGAAGACGCCCTTATACAAACGTCAAAA 2462
Qy      538 AlaThrTyrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSer 557
Db      2463 GCAACATATTCGTTTACGCTCAAGCTGTAAACCTTAAATATATCATTTGTTTGCTCA 2522
Qy      558 AaspAlaProAlaSerThrThrAlaSerValThrLeuValAenAaspLeuAenValIle 577
Db      2523 GATGACACAGTAGACAGACGCGCATCACTTAAGTAATGAATTAAGACTTAAGTATAC 2582
Qy      578 ThrAlaProAenGlyThrArgTyrValGlyAenAaspPheSerAlaProPheAenAen 597
Db      2583 ACTGCACCAATGGAACCTTAATCGTCGGAATGACTTTACGACACCGTATATTAACAAT 2642
Qy      598 TrpAaspGlyAArgAenAenValGluAenValPheIleAenSerProGlnSerGlyThrTyr 617
Db      2643 TGGGATGCGAAGAAACAAGTGAATATGTTATCAATGCTCTCAAGCGGACGATAT 2702
Qy      618 ThrIleGluValGlnAlaTyrAenValProValGlyProGlnAenPheSerLeuAlaIle 637
Db      2703 ACAAGTGAAGTCAAGCTTACATGATCAAGTAAGTCCGCAAACTTTCTTTAGCGATT 2762
Qy      638 ValAen 639
Db      2763 GTACAT 2768

```

RESULT 13
 LOCUS BD062155 3003 bp DNA linear PAT 27-AUG-2002
 DEFINITION Nucleic acids encoding a polypeptide having protease activity.
 ACCESSION BD062155
 VERSION BD062155.1 GI:22607760
 KEYWORDS JP 2001514529-A/39.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 3003)
 REFERENCE Sloana A. and Christianson, L.
 Nucleic acids encoding a polypeptide having protease activity

JOURNAL Patient: JP 2001514529-A 39 11-SEP-2001;
COMMENT NOVO NORDISK BIOTECH INC
PN JP 2001514529-A/39
PD 11-SEP-2001
PF 09-JUN-1998 JP 1999503145
PR 12-JUN-1997 US 08/873479
PI ALAN SIOWA, LYNN CHRISTIANSON
PC C12N15/57, C12N15/75, C12N9/54, C12K14/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key: Location/Qualifiers.
FEATURES
source 1..3003 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,75e-143 Length: 3003
Score: 2799.00 Matches: 541
Percent Similarity: 92.44 Conservative: 52
Best Local Similarity: 84.34 Mismatches: 45
Query Match: 85.74 Indels: 4
Gaps: 2
DB: 6
US-10-784-870-4 (1-639) x BD062155 (1-3003)
QY 1 Metatgls-----LyelyvalPheleuserValleuserAlaAlaIleleu 17
DB 846 ATGAGAAAGAAAGATCGAAGAGGTTTTTTATCGTTTATCGTTGCTGCATATTG 905
QY 18 SerThrValAlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
DB 906 TCTTCTGTGCTTAAAGAGCTCTTCTACTATTGGGCGGCAACATTTTGAATTTGACCTTT 965
QY 38 LysGlyIleGlnThrThrAspValSerGlyPheSerGlyGlnArgGlnThrGlyAla 57
DB 966 AAGGGATAGAGACCTTACGCTAGAGAAAGCTGCCACCAAGCAAGAAAGAAAGCGGAAAG 1025
QY 58 AlaAlaPheLeuValGlySerGlyAsnValIleLeuLeuGlyLeuLeuIleValLeu 77
DB 1026 GCATCTTTCTTGTAACTGTAAGATGTGAATAATCCCAAGAGATTCAAAGAAACTA 1085
QY 78 GluThrValProAlaAsnAsnIleValIleValGlnPheAsnGlyProIleLeuGlu 97
DB 1086 GAAGTAGTTCCAGCGGATACAAAGCTATATATGTTCAATTTGAGCGGACTATTTTAAAG 1145
QY 98 GluThrLysGlnIleLeuGluThrThrGlyAlaIleValIleLeuAspThrIleProAspThr 117
DB 1146 GAAGCGCACTTCAACTAGAGAGAGCGGAGGAAATTTCTGATTACATACAGATTAC 1205
QY 118 AlaThrIleValGlyIleGluGlyIleValAspValGlnSerIleValArgSerIleGluIleVal 137
DB 1206 GCTTATATGTCGATATGATGATGGATGTAAAGCGCTAACTAAAGCAATTCGCAATTGG 1265
QY 138 GluSerValGluProThrLeuProIleThrIleAspProGlnLeuPheThrIleGly 157
DB 1266 GAATCGGTTGAACCATATTTACTTTATATAAATGAGCCCGCAATTTTTCAGAGAGA 1325
QY 158 AlaSerThrLeuValIleAlaLeuAlaLeuAspThrLysGlnAsnAsnIleGlyValGln 177
DB 1336 GCTTCTGAATTATGTAACAGTACCTTACATTAAGCAAGAAAGATTAAGATACGT 1385
QY 178 LeuArgGlyIleGluGluIleAlaGlnIleValAlaSerAsnAspValHisThrIleThr 197
DB 1386 TTAAGAGATTGGAAACAATTGCCCAATACCGGCAAAATTAATGATTTTATACGTAAAC 1445
QY 198 AlaLysProGluThrLysValIleMetAsnAspValAlaArgGlyIleValIleAlaAspVal 217
DB 1446 CCAAGGCTGTAATACAGAGTTTGAATGACGTGGCCCGTGGCATGTGTAAGACACAGCTC 1505
QY 218 AlaGlnSerSerThrGlyLeuThrGlyGlnGlyIleValAlaValAlaAspThrGly 237

DB 1506 GCACAAATTAACCTTGGCTTATATGACAGACAGATTTGACAGTTCCTGATATCTGG 1565
QY 238 LeuAspThrGlyArgAsnAspSerSerMetHisGlyAlaPheArgGlyLysIleThrAla 257
DB 1566 CTTGATACAGAGAAATATGACATTCATGATGATGATGATCCCGGATTAATTAACCGCA 1625
QY 258 LeuThrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisVal 277
DB 1626 CTATATGACCTGGGCGAGAACCAATTAACGCCAATATCAATTAAGACATGAAACCAAGTT 1685
QY 278 AlaGlySerValIleGlyLysGlyValaThrAsnIleGlyMetAlaProGlnAlaAsnLeu 297
DB 1686 GCTGATCTGTGTTAGGAAAT---GCTACAAATTAAGGATGGCAACCGCAAGCCAACTTA 1742
QY 298 ValPheGlnSerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGln 317
DB 1743 GTCTTTCAATCTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
QY 318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerThrGlyAla 337
DB 1803 ACATTAATTCAGTCAAGCATATATGCTGGAGCGGAATTCATTAAGAAATTCATGGGGGCT 1862
QY 338 AlaValAsnGlyAlaThrThrThrAspSerArgAsnValaAspAspThrValArgLysAsn 357
DB 1863 CCAATTAACGATGCTTATACACAGACTCTGAAATGTTGATGATTTATGTAGAAAAAT 1922
QY 358 AspMetThrIleLeuPheAlaIleGlyAsnGlyAsnGlyProAsnGlyThrIleSerAla 377
DB 1923 GATATGACGATTTCTTTTGGCGCGGAATATGAGGACCAAGTACCGGTAACAATCAGTCA 1982
QY 378 ProGlyThrAlaLysAsnAlaIleThrValGlyValaThrGluAsnLeuAspProSerPhe 397
DB 1983 CCAGAACACAGAAAAAATGCAATTAACATGATGGGCAACCGAAACCTACGCTCAAGCTTC 2042
QY 398 GlySerThrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLys 417
DB 2043 GGATCTTATGCGGATTAATATTAACATGTTCTCTCAATTCCTTCAAGAGCTCTTAAGA 2102
QY 418 AspGlyArgIleLysProAspValMetAlaProGlyThrThrIleLeuSerAlaArgSer 437
DB 2103 GATGACGATATTAAGCGGAGCTCATGGCACAGTACGATATTTCTCTGCTAGATCA 2162
QY 438 SerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysThrAlaThrMetGly 457
DB 2163 TCATTAAGCTCCAGATTCCTCAATTCGGGCAAAACCATGATATTAATATGCTTACATGGGT 2222
QY 458 GlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPhe 477
DB 2223 GGTACTTCTATGCTACTCCCAATTTGACAGGTATGTCACAAATTAAGGAGCATTTT 2282
QY 478 ValLysAsnArgGlyIleThrProLysProSerLeuLeuValaLeuIleAlaGly 497
DB 2283 GTGAAATAATAGAGGGTAACCTTAAGGCTTCCTTTTAAAGCTGTTAATTTGCAAGT 2342
QY 498 AlaAlaAspValGlyLeuGlyThrProAsnGlyAsnGlnGlyThrProGlyArgValaThrLeu 517
DB 2343 GCTCGGATGTGGACCTGCTTCCAAATGTGAACCAAGATGGGGAAGATGACGTGA 2402
QY 518 AspLysSerLeuAsnValAlaThrValaAsnGlySerSerAlaLeuSerThrSerGlnLys 537
DB 2403 GATTAATCCCTTAATATGTCGATTTGTGAATGAACGAGCCCTTTATCAACAGTCAAAA 2462
QY 538 AlaThrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValaIlePheSer 557
DB 2463 GCAACATATTCGTTTAAAGGCTCAAGCTGGTAAACCTTTAAATAATCACTGTTGGTCA 2522
QY 558 AspAlaProAlaSerThrThrAlaSerValaThrLeuValaAsnAspLeuAspLeuValaIle 577
DB 2523 GATCAACAGGTATACAGACGCGATCACTTAATTTAGATGATTTAGATCTTGATATC 2582
QY 578 ThrAlaProAsnGlyThrArgThrValaGlyAsnAspPheSerAlaProPheAspAsn 597

Db	2583	ACTGCACGAAATGGAACCTAAATACGTCCGAAATGACTTTACAGCACCGTATGATACCAT	2642
Qy	598	TTpaaaglyAaAaAaAaAaValGluAaAaValPheIIeaAaSerProGlnSerGlyThrTy	617
Db	2643	TGGATATGCGAAGAAACACCTGAGAAATGCTTATCAATGCTCTCAAGCGAAGCAT	2702
Qy	618	ThrlleGluValGlnAblleThyAaAaValProValGlyProGlnAaAaPheSerleuAlaIle	637
Db	2703	AAAGTGAAGTCAAGCTTACCAATGATACCAAGTAAAGTCCGAAACCTTTCTTAAAGCAT	2762
Qy	638	Valaen 639	
Db	2763	GTACAT 2768	
RESULT 14			
LOCUS	E03808	2218 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA encoding alkaline protease Ya.		
ACCESSION	E03808		
VERSION	E03808.1	GI:2172022	
KEYWORDS	JP 1992197182-A/1.		
SOURCE	Bacillus sp.		
ORGANISM	Bacillus sp.		
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
AUTHORS	1 (bases 1 to 2218)		
TITLE	Iobe,S., Odera,M. and Aaai,Y.		
JOURNAL	DNA CODING ALKALINE PROTEASE YA ENZYME AND PRODUCTION OF ALKALINE PROTEASE YA USING THE DNA Patent: JP 1992197182-A 1 16-JUL-1992;		
COMMENT	LION CORP		
OS	Bacillus sp.		
PN	JP 1992197182-A/1		
PD	16-JUL-1992		
PF	28-NOV-1990 JP 1990327110		
PI	TOBE SRIICHT, ODBA MOTOMASU, ASAI YOSHIO		
PC	C12N15/57, C12D1/386, C12N9/54, (C12N15/57, C12R1:07), (C12N9/54,		
PC	C12R1:07);		
CC	strandedness: Double;		
CC	topology: linear;		
CC	hypotheical: No;		
CC	anti-sense: No;		
CC	*source: strain=Y strain;		
CC	*source: clone=pUB8A;		
EH	Key Location/Qualifiers		
FT	CDS 218..2125		
FT	1. .2218		
FT	/product='precursor of alkaline protease Ya'		
FT	sig_peptide 218..823		
FT	mat_peptide 824..2122		
FT	/product='Precursor of alkaline protease Ya'.		
FEATURES	Location/Qualifiers		
source	1..2218		
ORIGIN	/organism="Bacillus sp."		
	/mol_type="genomic DNA"		
	/db_xref="taxon:1409"		
Alignment Scores:			
Pred. No.:	2.27e-127	Length:	2218
Score:	2500.50	Matches:	480
Percent Simlarity:	87.9%	Conservative:	80
Best Local Simlarity:	75.4%	Mismatches:	72
Query Match:	76.5%	Indels:	5
DB:	6	Gaps:	4
US-10-784-870-4 (1-639) x E03808 (1-2218)			
Qy	3	LyalsyAaValPheIleuSerValIleuSerAlaAlaIleuSerThrValAlaIleu	22
Db	227	AAAAAAGAGAGAGCGTATCAGTAGTGCTCCCGCTCGAATCTTAAGCTCAGTTATGGTT	266
Qy	23	AaAaAaProSerAlaGlyAaAaAaThrPheAaPheAaAaPheIleuSerGlyIleGlnThr	42

Db	287	AGTTACCACTAGTGGG-----GCAGATTTTCAATGAAATTTTAATGCTGAAAGT	340
Qy	43	ThrlAspValSerGlyPheSerIysGlnThrGlyAlaIleAlaPheLeuVal	62
Db	341	TTAGAAAAGCTAGT---TTGGTTAAACCGATTAAGTACGGGTACAGGACCTTTCTTAGTA	397
Qy	63	GlusSerGluAsnValLeuLeuLeuIleuIysGlyLeuLeuValLeuGluThrValProIa	82
Db	398	GATACGGAAAATATTAATATCTTCAAGGATTTCAAAAAGAGCTAAACAGACAGACGAAG	457
Qy	83	AsnAspIleValIleValGlnPheAsnGlyProIleLeuGluGluThrIysGlnIys	102
Db	458	GATTAACGAACCTTAACCTGTAACAATTTACTGCAACAATTTCAAGGAAAGAGCAAAAAGA	517
Qy	103	LeuGluThrThrGlyValAlaValIleLeuAspTyrIleProAspTyrAlaTyrIleValGlu	122
Db	518	TTAGAGTCTCTAGAGATACGATTTCTAGTTATAGTTCCAGATTAATGCTTTTATGTTCCAG	577
Qy	123	TyrGluGlyAspValGlnSerIleValArgSerIleGluIleValGluSerValGluPro	142
Db	578	TATAGTGGGTACAAAATAAT---ATAAGTACTTAATCAATCTGTTGAGAAACGTACAAACA	634
Qy	143	TyrLeuProLysTyrIleValIleAspProGlnLeuPheThrIleGlyValAspThrIleVal	162
Db	635	TTTTTACCAATTATATTAATTAATGATCTCTGCTTTTAAACGAAGGTCTTCCAGCTTGTT	694
Qy	163	LybAlaLeuAlaLeuAspThrIleValGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyIleGlu	182
Db	655	CAAGGGGTATTTTAAATATACAAACACGAAATATTAACATGAATTTACCGGTTTAGAT	754
Qy	183	GluIleAlaGlnTyrValAlaSerAsnAspValIleTyrIleThrAlaIleAspProGluTyr	202
Db	755	GAGATCGTTACATATGCTGCAAAATATATGATGTCTTATATATCAACAAAGCCGAGTAT	814
Qy	203	LybValMetAsnAspValAlaArgGlyIleValLybValAspValAlaGlnSerSerTyr	222
Db	815	GAGCTAAAGAATGAATGACGACGAAGGAGTATGTAATTAAGCTGAATGTGCACAAAACATTAC	874
Qy	223	GlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArg	242
Db	875	GGATTATATGACAAAGTCAACTAGTATGCAATGAGGACACAGGCTTATAGTACAGGCTCT	934
Qy	243	AsnAspSerSerMetIleGlyAlaPheAspArgGlyLysIleThrAlaLeuTyrAlaLeuGly	262
Db	935	AACGATATCTTCTATGATCAAGAGCATTCGCGCGGAAATCAACAGCTCTTTACCGCTTAGGA	994
Qy	263	ArgThrAsnAsnAlaAsnAspThrAsnGlyIleGlyIleThrIleValAlaGlySerValLeu	282
Db	995	AGAACTAATTAATGCAAGTGAATCCGAATGGGACATGGACACATGTAGCAGGTTCTTAGCTT	105
Qy	283	GlyAsnGlyAlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIle	302
Db	1055	GGTAT---GCTTAAATAAAGAAATGAGCTCCGCAAGCTAATTAAGTCTTCAACCTAAT	1111
Qy	303	MetAspSerSerGlyGlyLeuGlyGlyLeuPheProSerAsnLeuGlnThrIlePheSerGln	322
Db	1112	ATGATAGCAACCGAGAGATTAGTGCGCTTACATCGAATCTTAATTAACCTTATTTAGTCA	1177
Qy	323	AlaPheSerAlaGlyAlaArgIleIleThrAsnSerTrpGlyAlaAlaValaAsnGlyAla	342
Db	1172	GCTTGGAATGCTGGAGCAAGAAATCATGACTAATCTTTGGGAGCCCAAGTAATGAGAGCG	1233
Qy	343	TyrThrThrAspSerAspAsnValAspAspTyrValaArgLybAsnAspMetThrIleLeu	362
Db	1232	TACACTGCTAACTGAGACAAAGTGAATGAATGTTCGAAATTAATGAATATGACGGTACTT	1299
Qy	363	PheAlaAlaGlyAsnGluArgProAsnGlyIleThrIleSerAlaProGlyThrAlaLys	382
Db	1292	TTTTCAGCTGTTATGAAAGCTCTTAATTCAGAAACAATTAATGCTTCCAGGTACACGCAAA	1355
Qy	383	AsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAsp	402
Db	1352	AATCTATTAATGAGTCCGGGCAACGAAACATATGCGCCAAACCTTCGTTGCATATGACAT	1411

Oy	403	Asn1IeAsn1IeVal1AaGlnPheSerSerArgGlyProThrIysAspGlyArg1Ile	422
Db	1412	AAACCAATATCATATTGACACATTTTTCATCGAGAGAGCTACGAGGATGACGAATTAAG	1471
Oy	423	ProAspValMetAlaProGlyThrThrIleLeuSerAlaArgSerSerLeuAlaProAsp	442
Db	1472	CCTACGCTAAACAGCTCCTCGAAACATTTATTTATACAGACGTTCTCTTGCTCGACAC	1531
Oy	443	SerSerPheThrAlaAsnIleAspSerIysThrAlaTyrrMetGlyGlyThrSerMetAla	462
Db	1532	TCTTCGTTTGGGCAATATATACAGTAATACCGGTATATGAGGGGGTACTCCATGGCG	1591
Oy	463	ThrProIleVal1AaGlyAsnVal1AaGlnLeuArgGlyIleIlePheVal1AsnAspGly	482
Db	1592	ACACCTATTTGTCAGAGGAATGTCGCCGAATTACGTAGCATTTTATATAAATAATAGGT	1651
Oy	483	IleThrProIysProSerLeuLeuVal1AaIleuIleAlaGlyVal1AaAspVal1Gly	502
Db	1652	ATTACTCTTAAGCCTCTTTATTAATAAGCTCACTTATCGCTGGCTACACGAATGGGT	1711
Oy	503	LeuGlyThrProAsnGlyAsnGlnGlyThrGlyArgVal1ThrIleAspLysSerLeuAsn	522
Db	1712	TTAGAGATATCTTAGAGTGACCAAGCGCTGGGGCGGTACTCTAGATAAATCGTTAAT	1771
Oy	523	Val1AlaTyrrVal1AsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrThrPhe	542
Db	1772	GTAAGGTATGTCANAGAAACAGCATTTAGCCACAGACAAATAACGATATCGTTC	1831
Oy	543	ThrAlaGlnThrAlaGlyLysProLeuLysIleSerLeuVal1TyrSerAspAlaProAlaSer	562
Db	1832	CAAGCACAAAGCGGGTAAACCTTTAAATACTCGTTAAGTATGACAGCATGCTCCGGAAGT	1891
Oy	563	ThrThrAlaSerVal1ThrLeuVal1AsnAspLeuAspVal1IleThrAlaProAsnGly	582
Db	1892	ACAACTGCATCTTATACCTAGTTATGATTTAATCTAGTTATTACTGCTCCGAATGGA	1951
Oy	583	ThrArgTyrrVal1GlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArgAsn	602
Db	1952	CAAAAATATGTAGAGAAATGATTTTATGTTATCTTATGATTAATTAACCTGGGATGCTGCAC	2011
Oy	603	AsnVal1GlnAsnVal1PheIleAsnSerProGlnSerGlyThrThrIleGlnVal1Gln	622
Db	2012	AATGTTGGAACGTATTTATTAACCTCCGCAATCTGGAACGTATATTAATGAGAGTTCAA	2071
Oy	623	AlaTyrrAsnVal1ProVal1GlyProGlnAsnPheSerLeuAlaIleVal1Asn	639
Db	2072	GCGTATATATGTCATCTGCGCCACAGCGTTTCTCATAGCTATGTCAT	2122
RESULT 15			
LOCUS	AX839476	1305 bp	DNA linear PAT 15-DEC-2003
DEFINITION	Sequence 2 from Patent EP1347044.		
ACCESSION	AX839476		
VERSION	AX839476.1	GI:39922766	
KEYWORDS			
SOURCE			
ORGANISM	Bacillus sp. KSM-KP43		
REFERENCE	Bacillus sp. KSM-KP43		
AUTHORS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
TITLE	Okuda,M.K., Sato,T.K., Sato,K.K., Sumitomo,N.K., Izawa,Y.K.,		
JOURNAL	Saeki,K.K., Kobayashi,T.K. and Nomura,M.K.		
	Alkaline protease		
	Patent: EP 1347044-A 2 24-SEP-2003;		
	Kao Corporation (JP)		
FEATURES	Location/Qualifiers		
source	1..1305		
	/organism="Bacillus sp. KSM-KP43"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:109322"		
	<1..1305		
	/note="unnamed protein product"		
	/codon_state=1		

	/transl_table=1	
	/protein_id="CAE85146.1"	
	/db_xref="gi:39922767"	
	/translation="NDVARGIVKADVAQSSVTGLGGQGIVAVADNGLDTGRNDSMHI AFREKTALVALGRITNNANDTNGHGTHTVASVLGAGSTNKGMAPDANI VPOSIMNSGC GLGLGPSNLQTLPFOAQSAGARIHTNSWGAENVAYTTDSNRVDYVRKNMTILPAA GENEPNGGTTIPSGATAKNALTVGATEENLRPFSGFVADNIINVAOFSRSPPTDGRIKI DVMAFGPTLILARSLSIAPDSFPFMANHDSELYVMGTSKATPIVAGNVQLOLEHFVKMI GITPKPSLLKALKIAGAADIGLVPNGOGKGRVTLDSILVANAYNESSISLTSGOKAI YSFTPATRAKEPLKISLWSMDAPASSTASTLVINDIDLVITAPNGIOYVANDPFSPPNDI WDGRNNVENVFINAPQSTYTILEVOAINVPGPOTFSIALYN"	
ORIGIN		
Alignment Scores:		
Pred. No.:	2,456-110	Length: 1305
Score:	2183.00	Matches: 118
Best Similarity:	99.3%	Conservative: 13
Best Local Similarity:	96.3%	Mismatches: 3
Query Match:	66.8%	Indels: 0
DB:	6	Gaps: 0
US-10-784-870-4 (1-639) x AM839476 (1-1305)		
Oy	206 AenapValAlaarGlylYleValYsaLaapValalaginSerTyrlYleuty	225
Dd	1 AATGATGTTGGCGCGTGGAATTGTCAAACGGATGTGGCTCAGACGAGCTAACGGTTGTAT	60
Oy	226 GlyInglYglnIlleValAlaValAlaapThrGlyLeuaBPTHrGlYArgaBnppSer	245
Dd	61 GGACAMGACAGATGTAAGCGGTTCGCATYACAGGSCYTGATACAGGTCCAAATACACT	120
Oy	246 SerMetHisglualaPhaeArgGlyLyseIleThraIaleutyralAleugIyarGThraen	265
Dd	121 TCGATGACAGAAGCCCTCCGCCGAAAATTACTGCATTATATCATTTGGACGACGAAAT	180
Oy	266 AsnaLaenaApPThraenGIYhiSgLYThrhIsValaagISeerValleugIyanSgly	285
Dd	101 AATGCCAATGANTAGAAATGTCATGTACATGTGGCTCGCTCATTTAGAAACGGC	240
Oy	286 AlaThraenLySgLYMeCaLabProGlnAlaaenLeuValPheGlnSerIleMetAepSer	305
Dd	241 TCACCTAATAAGAAATGCGCCTCAGGGCAATCAATGCTTCCAATCTATCATGTAGATAC	300
Oy	306 SerGlyGlyLeugIYGlyLeuProSerBnneugiImrrLeuPheSerGlnAlaPheSer	325
Dd	301 GGTGGGGGACTTGGAGACACTTCTGAATTCGCAACTTATTCAGCCAGACATACACT	360
Oy	326 AlaGlyAlaargIleHisThraBnserTPGlyalaalavalasngIyalatYrThTrhr	345
Dd	361 GCTGGTCCAGAAATTCATACAACCTCTGGGGAGCGACGATGAATGGGCTTACACAACA	420
Oy	346 AppSerArgaBnvalAspaBPTYrValargLYbaBnppMerThrIlleuPheaIaala	365
Dd	421 GATTCACAAAAATGTGATGATCATATGTGCGCAAAAATGAAATGATGACATCTTTTGGCTGCGC	480
Oy	366 GlyaBngIuaArgProBaBngIyglyYhrIleSeraIaaproGlyThraIalyBnaIalie	385
Dd	481 GGGAAATGAAGAACCGAAACGGCGAAACCAATCAAGTGCACAGGACACAGCTAAATAATGCAAT	540
Oy	386 ThrValaGlyAlaThrgIubBnneulyArgProSerPheGlySerTYrnlAaBnaIleaon	405
Dd	541 ACAGTCGAGACTACGGAAAACCTCCGCCAACCCTTGGGCTCTTAATGGGACAAATACAAC	600
Oy	406 HisValaIagInPheSerSerArgGlyProThrylaBpgIylarGlyleuyBProabval	425
Dd	601 CATGTGGCACAGATTCTTCACTGTCGACCGAACAAAGATGACGAGATCAAAACGAGATGTC	660
Oy	426 MetCalapProGlyThrTyrlleuSeraIaayGerSerleuAlaProabSerSerPhe	445
Dd	661 ATGGACACCGGAAACGTTCACTATCTACGAAAGATCTTCTTGGACCGGAATTCCTCTTC	720
Oy	446 TPraLaBnHisApsSerLyTyralaTyrmelGlylThrSerMetAlaThrProile	465

```
Db      721 TGGGCGAACCATTGACGTAATATGCAATCATGGGTGGAACGTTCATGGCTACACCGATC 780
Qy      466 ValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIlyAsnArgGlyIleThrPro 485
Db      781 GTTGTGGAACGTGGCAACGTTCTGGAGCATTTGTGAAAAACAGAGCATCACACCA 840
Qy      486 LysProSerLeuLeuLysAlaLeuIleAlaGlyValAlaAspValGlyLeuGlyTyr 505
Db      841 AGCCCTTCTCTATTAAAGCGGCACTGATGGCCGGTGCACCTGACATCGGCTTGGCTAC 900
Qy      506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db      901 CCGAACGGTAACCAAGATGGGAGCAGTGCATTTGAATCCCTGAACGTTGCCTAT 960
Qy      526 ValAsnGlySerSerLalaSerThrSerGlnLysAlaThrTyrThrPheThrAlaThr 545
Db      961 GTGAACGAGTCCAGTCTCTATCCACCAAGCAAAAGCGACGTACTCGTTACTGCTACT 1020
Qy      546 AlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAlaSerThrAla 565
Db      1021 GCCGGCAACCTTTGAAATCTCCCTGTAATGCTGATGCCCTGCGAGCACACTGCT 1080
Qy      566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db      1081 TCCGTAAACGCTTGTCAATGATCTGACCTTGTCAATACCGCTCAATGGCACACAGTAT 1140
Qy      586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyValArgAsnValGlu 605
Db      1141 GTAGGAATGACTTAACTTCGCCATCAATGATTAACCTGGGATGGCCGCAATAACTAGAA 1200
Qy      606 AsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGlnAlaTyrAsn 625
Db      1201 AATGATTTTATTAATCACCAACAGCGGACGTTATACATTAAGGTACAGGCTTATAAC 1260
Qy      626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db      1261 GTACCGGTGGACCAACAGACTTCTCGTTGGCAATTGTGAT 1302
```

Search completed: April 8, 2006, 05:26:55
Job time : 7515 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 03:09:11 ; Search time 1640 Seconds
(without alignments)
2596.792 Million cell updates/sec

Title: US-10-784-870-4
Perfect score: 3267
Sequence: 1 MKKKVFLSLVLSAAAILSTV.....EVOAVNVPGPNPSLATVN 639

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame.p2n.model -DEV=xlp
-O=/abs8/ABSSWEB-epool/US10784870/runat_07042006_14314_27447/app_query.fasta_1
-DB=N Geneseg -OFMT=faetap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs802p
-USER=US10784870@cgc1_1_1923@runat_07042006_14314_27447 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2004as.*
12: geneseqn2004bs.*
13: geneseqn2004cs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3267	100.0	1920	2	AAK37277
2	3072	94.0	1923	2	AAK37279
3	3066	93.8	1923	2	AAK37278
4	3066	93.8	1923	12	ADM40772

5	3066	93.8	1923	13	ADSL4428	AdSL4428 Bacillus
6	3066	93.8	1923	13	ADT49605	AdT49605 Bacillus
7	2799	85.7	3003	2	AAV82382	AAV82382 Bacillus
8	2183	66.8	1305	11	ADY33779	ADY33779 Bacillus
9	2183	66.8	1305	12	ADL25803	ADL25803 Bacillus
10	1990.5	60.9	1399	2	AAQ27516	AAQ27516 Alkali-pr
11	477	14.6	1977	2	AAV85667	AAV85667 Thermococ
12	477	14.6	1977	2	AAV85692	AAV85692 M9856926
13	443	13.6	1962	2	AAV85695	AAV85695 Pyrococcus
14	443	13.6	1962	2	AAV85929	AAV85929 Hyperther
15	443	13.6	1962	2	AAV85929	AAV85929 Hyperther
16	435	13.3	110000	12	ADN46845_14	ADN46845_14
17	435	13.3	110000	12	ADN47591_05	ADN47591_05
18	435	13.3	110000	12	ADN47591_06	ADN47591_06
19	435	13.3	110000	12	ADN46123_14	ADN46123_14
20	435	13.3	110000	12	ADN47209_05	ADN47209_05
21	435	13.3	110000	12	ADN47209_06	ADN47209_06
22	435	13.3	110000	12	ADN46464_14	ADN46464_14
23	435	13.3	110000	12	ADN47960_05	ADN47960_05
24	435	13.3	110000	12	ADN47960_06	ADN47960_06
25	432	13.2	1977	2	AAV85669	AAV85669
26	417.5	12.8	1336	2	AAV85920	AAV85920
27	417.5	12.8	1336	2	AAV85668	AAV85668
28	392	12.0	2121	4	ABL54900	ABL54900 T. yonsei
29	382	11.7	3684	12	ADM99122	ADM99122 Bacterial
30	357.5	10.9	2539	2	AAV61454	AAV61454
31	357.5	10.9	2809	2	AAV61455	AAV61455 Dnpa-mel
32	345.5	10.6	1656	12	ADM99240	ADM99240 Bnvirome
33	344.5	10.5	135638	10	ABX34289	ABX34289 S. atrool
34	343.5	10.5	104096	13	ADX56092	ADX56092 Streptococ
35	342.5	10.5	1329	6	ABK74643	ABK74643 Bacillus
36	339.5	10.4	1332	12	ADM99124	ADM99124 Bacterial
37	326	10.0	2835	2	AAV88141	AAV88141 Hyperther
38	326	10.0	4765	2	AAV88132	AAV88132 Pyrococcus
39	326	10.0	4765	2	AAV85670	AAV85670 Pyrococcus
40	326	10.0	4765	2	AAV85921	AAV85921 M9856926
41	325.5	10.0	89378	12	ADN47591_20	ADN47591_20
42	325.5	10.0	89378	12	ADN47209_20	ADN47209_20
43	325.5	10.0	89378	12	ADN47960_20	ADN47960_20
44	325.5	10.0	110000	12	ADN46845_00	ADN46845 Thermococ
45	325.5	10.0	110000	12	ADN46123_00	ADN46123 Thermococ

ALIGNMENTS

RESULT 1	AAK37277	AAK37277 standard; DNA; 1920 BP.
ID	AAK37277	AAK37277 standard; DNA; 1920 BP.
XX	AAK37277	
AC	AAK37277	
XX	AAK37277	
DT	20-MAR-2003	(revlsed)
DT	21-JUL-1999	(first entry)
XX	XX	
DE	Bacillus alkaline protease encoding DNA.	
XX	XX	
KW	Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;	
KW	washing composition; oxidising agent; ss.	
XX	XX	
OS	Bacillus sp.	
XX	XX	
PN	W09918218-A1.	
XX	XX	
PD	15-APR-1999.	
XX	XX	
PF	07-OCT-1998;	98WO-JP004528.
XX	XX	
PR	07-OCT-1997;	97JP-00274570.
XX	XX	
PA	(KAO) KAO CORP.	
XX	XX	
PI	Takaiwa M, Okuda M, Sasaki K, Kubota H, Hitomi J, Kageyama Y;	
PI	Shikata S, Nomura M,	

XX WPI: 1999-287736/27.
DR P-PSDB; AAY17087, AAY17089.
XX Alkali protease from *Bacillus* used in washing powders.
XX
PS Disclosure: Page 53-58; 71pp; Japanese.
XX The invention relates to alkaline proteases produced by strains of
CC *Bacillus*. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1920 BP; 576 A; 411 C; 473 G; 460 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,736-243 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-784-870-4 (1-639) x AAX37277 (1-1920)
QY 1 MetAglYsLySvLyVAlPheLeuSeRValLeuSeRAlaAlaIleuSeRThrVal 20
DB 1 ATGAGAAAGAAAGAGAGTGTGTTTATCTGTTTATCACTGCGACGATCTGTGCACTGTT 60
QY 21 AlaLeuAenAenProSeRAlaGlYAspAlaArgThrPheAspLeuAspPheLySgIlyle 40
DB 61 GCATTAAACAAATCCCTCGGCTGTGATGCAAGACATTGTTGATCTGATTTTAAAGAAAT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerLySgInaArgGlnThrGlyAlaAlaAlaPhe 60
DB 61 GCATTAAACAAATCCCTCGGCTGTGATGCAAGACATTGTTGATCTGATTTTAAAGAAAT 120
QY 41 GlnThrThrThrAspValSerGlyPheSerLySgInaArgGlnThrGlyAlaAlaAlaPhe 60
DB 121 CAAGAAACAAACGAGATGTCAGTGTGTTCTCCAAACGCGACAAACAGGTCGCTCATTT 180
QY 61 LeuValIGluSeRGIAsnValIlySLeuLeuLySgIlyLeuLeuLySLeuGluThrVal 80
DB 181 CTGGTGGAGCTCGAAATGTGAACCTTCTTAAAGATTCCTTAAAGAACTTGAAACAGTA 240
QY 81 ProAlaenAenLySLeuHileValGlnPheAenGlyProIleLeuGluGluThrLyS 100
DB 241 CCGGCAAAATTAATTAACCTCAATATGTCCAATTCATATGGCCCCCATTTTGAAGAAACAAA 300
QY 101 GlnLySLeuGluThrThrGlyAlaLySileuApyrTlleProAspTyrAlaTyrTlle 120
DB 301 CAGAAGCTAGAGAACAACTGAGCAAGATTCCTGACTACATCCCTGATTTATGCAATATTT 360
QY 121 ValGluYrGluGluYAspValGlnSerLySValArgSerIleGluHileValGluSerVal 140
DB 361 GTCCAGTATGAGGGGAGATGTTCAATCAAAAGTCCGCTCATTTGAAACCTGGAATCAAGT 420
QY 141 GluProTyrLeuProLySLeuTyrLySileAspProGlnLeuPheThrLySgIlyAlaSerThr 160
DB 421 GAGCCATATCTTCCGAAATATCAAAATATAGATCCCGAGCTTTTCAAAAAGCGCATCGACG 480
QY 161 LeuValIySAlaLeuAlaLeuAspThrLySgInaenAenLySgInaValGlnLeuArgGly 180
DB 481 CTGGGAAAGCGTTGGCTTGATCGAAGCAAAACATTAAGATGCAATTAAGAGAGC 540
QY 181 IleGluIleIleAlaGlnTyrValAlaSerAenAspValHileTyrIleThrAlaLySPro 200

DB 541 ATCGAGAAATCGCTCACTAGTACAGAACATGACATTCATTAATTAACGCAAGCCT 600
QY 201 GluTyrLySValMetAenAspValAlaArgGlyIleValIlySAlaAspValAlaGlnSer 220
DB 601 GAATTAAGAGTGAATGAATGATGTGCCAGAGGTATTTGCAAAAGCGAGTGTGCACAGAGC 660
QY 221 SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
DB 661 AGCTACGGTGTGTATGACAAAGCCAGATTTGTCCAGTTGCCGATCTGGAATGGATACA 720
QY 241 GlyArgAenAenSerSerMetHileGlyAlaPheArgGlyLySileThrAlaLeuTyrAla 260
DB 721 GGAAGAAACGACAGTTCATGCAATGCAATGACCTTCCCGGTTAAATTAACGACATTAATCA 780
QY 261 LeuGlyArgThrAenAenAlaAsnAspThrAenGlyHileGlyThrHileValAlaGlySer 280
DB 781 CTGGGTCCGAGCAATTAATGCAATGATGCAACGTCATGTTACCCAGTGGCAGGTTCG 840
QY 281 ValLeuGlyAenGlyValaThrAenLySgIlyMetAlaProGlnAlaAsnLeuValPheGln 300
DB 841 GTATTAGGAATGGCGCAACCAATTAAGAAATGGCACTCAAGCGAATCTGATTTTCA 900
QY 301 SerIleMetAenSerSerGlyLySLeuGlyLySLeuProSerAenLeuGlnThrLeuPhe 320
DB 901 TCCATCAATGATATGACAGTGTGGGCTTGAGGCTTGCTTCATCTGCAACCTTATTC 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHileThrAenSerTyrGlyAlaAlaValAen 340
DB 961 AGCCAAAGCATTCAGTGCAGTGCAGCAATTCATCAAACTCTGGGGGCGAGCGTGAAT 1020
QY 341 GlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLySAsnAspMetThr 360
DB 1021 GGGGCTACAGACAGATTCAGAAATGTGAATGATGTAATGTAAGGAAATGAATATGACG 1080
QY 361 IleLeuPheAlaAlaGlyAenGluArgProAenGlyLySThrIleSerAlaProGlyThr 380
DB 1081 ATTCTTTTGGGGCTGGGAATGAAGCGCGAAGCGGCTTACATCAATGACACTGGTATCG 1140
QY 381 AlaIySAsnAlaIleThrValGlyAlaThrGluAenLeuAArgProSerPheGlySerTyr 400
DB 1141 GCTAAACCGCATTAACAGTGTGCGCGCAACCAAACTCGCTCCAAAGCTTGGTCTCAT 1200
QY 401 AlaAspAenIleAenHileValAlaGlnPheSerSerArgLySProThrLySAspGlyArg 420
DB 1201 GCAATTAATTAATTAACAGTGTGCAAGTTCCTTCCGTGGCGCGCAAAAGATGGGCGA 1260
QY 421 IleLySProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAAGCTGATGTGATGGCGCGCAAGGACATTAATTAACAGAAAGATCTTCTTCA 1320
QY 441 ProAspSerSerPheTyrAlaAsnHileAspSerLySValaTyrMetGlyLySThrSer 460
DB 1321 CCGCATTCCTCTTCTGGCGCAATCATGACGAAATATGCTTATATGGGTGGAACGTCC 1380
QY 461 MetAlaThrProIleValAlaGlyAenValAlaGlnLeuAArgGluHilePheValIySAsn 480
DB 1381 ATGGCAACAGGATGTGTGGGGAATGTGACAGCTCCGTGAGCATTTTGTAATAAT 1440
QY 481 ArgGlyIleThrProLySProSerLeuLeuLySAlaAlaLeuIleAlaGlyAlaAlaAsp 500
DB 1441 AGAGGAATCACTCTTAACCTTCCCTATTTAAAGAGCTTTGATTTGCAAGGCGCTGAT 1500
QY 501 ValGlyLeuGlyTyrProAenGlyAenGlnGlyTyrGlyArgValThrLeuAspLySAsp 520
DB 1501 GTTGATTTGGGTATATCCCAACGAAACCAAGATGGGGCCAGATGACCTGATTAATCG 1560
QY 521 LeuAenValAlaTyrValAenGlySerSerAlaLeuSerThrSerGlnLySAlaThrTyr 540
DB 1561 TTGAACGTGTGCTATGTAAGAAATCCAGTGCCTTATCACTGACCAAAAGCGACATAT 1620
QY 541 ThrPheThrAlaThrAlaGlyLySProLeuLySileSerLeuValTyrSerAspAlaPro 560
DB 1621 ACCTTTATGCAAGCGGGGCAAGCATTTGAATATCTCCCTGGTATGTTCCGATGCCCT 1680

QY 561 AAlaSerThrThraLeuSerValThrLeuValAlaAspLeuLeuValIleThrAlaPro 580
 DB 1681 GGAAGACACTGACTGCTCTGTGTAACCTGGTCAGTAATTTGGATTGCTATACAGACCA 1740
 QY 581 AaagIyThrArgTyValGlyAsnAspPheSerAlaProPheAspAsnIleTPAPGly 600
 DB 1741 AACGGAAACAAGATATGTCGGGAATGACTTCTCAGCACATTTGACATTAACCTGGATGGC 1800
 QY 601 ArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTrpIleGlu 620
 DB 1801 CCCAATTAACGTGAAAATGATATTATTAATTCGCCCAAGTGAACATATACCATTTGAG 1860
 QY 621 ValGlnAlaLeuThrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
 DB 1861 GTGCAAGCATATATATGTGCGGTGGACCAACAACTTCTGTCGCAATTTGAAAC 1917
 RESULT 2
 AAX37279
 ID AAX37279 standard; DNA; 1923 BP.
 AC AAX37279;
 XX 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 XX
 DE Bacillus alkaline protease encoding DNA.
 XX
 KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KM washing composition; oxidising agent; ss.
 XX
 OS Bacillus sp.
 XX
 FN WO9118218-A1.
 PD 15-APR-1999.
 XX
 PP 07-OCT-1998; Y98WO-UP004528.
 XX
 PR 07-OCT-1997; 97JP-00274570.
 XX
 PA (KAO) KAO CORP.
 PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 PI Shikata S, Nomura M;
 XX
 DR WPI; 1999-287736/27.
 XX
 DR P-PsDB; AAV17091.
 XX
 PT Alkali protease from Bacillus used in washing powders.
 PS
 PS Disclosure; Page 63-68; 71pp; Japanese.
 XX
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease encoding
 CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
 XX
 SQ Sequence 1923 BP; 578 A; 417 C; 474 G; 454 T; 0 U; 0 Other;

Alignment Scores: 4,62e-228 Length: 1923
 Score: 3072.00 Matches: 596

Percent Similarity: 97.0% Conservative: 23
 Best Local Similarity: 93.4% Mismatches: 19
 Query Match: 94.0% Indels: 0
 DB: 2 Gaps: 0
 US-10-784-870-4 (1-639) x AAX37279 (1-1923)
 QY 2 ArgGlyValGlyValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
 DB 7 AAGAAAGAAAAGGTGTTTATCTGTTTATCTGATGCGCAGCCATTTTTCGACTGTGGG 66
 QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIleGlu 41
 DB 67 TTAAGTAACTCATCTGCGAGGTGTGAGAAATTTGAATCTGAATTCAGAGAAATTCAG 126
 QY 42 ThrThrThrAspValSerGlyPheSerIyGlnArgGlnThrGlyValAlaAlaPheLeu 61
 DB 127 ACAACAACAGATGCTAAAGGTTTCTCCAGACAGGGGCAAGCTGTCTCTCTTTCTG 186
 QY 62 ValGlnSerGluAsnValIleLeuLeuIyGlyLeuLeuIyGlyLeuGluThrValPro 81
 DB 187 GTGGAATCTGAAAATGTGAACTCCCAAAAGGTTTGCAGAAAGAGCTTGAACAGTCCCG 246
 QY 82 AlaAsnAsnLeuLeuHieIleValGlnPheAsnGlyProIleLeuGlnGluThrIyGln 101
 DB 247 GCAATTAATTAATCTCATATTAATCAATTCAGATGACCAATTTTGAAGAAACAAACAG 306
 QY 102 IyLeuGluThrThrGlyValAlaIySileuAspTyIleProAspTyAlaTyIleVal 121
 DB 307 CAGCTGGAAAAACAGGGGCAAAAGTTCTCGACTACATACCTGATTATGCTTACATTGTC 366
 QY 122 GluTyGluGlyAspValGlnSerIyValArgSerIleGluHieIleValGlnSerValGlu 141
 DB 367 GAGTATGAGGGCGGATGTTAAGTCAGCAACAGCACATTCAGACAGCTGGAAATCCGTGGAG 426
 QY 142 ProTyIleuProIyTyIleSileuAspProGlnIleuPheThrIyGlyValAspSerThrLeu 161
 DB 427 CCTTATTTGCGCATTAATACGAATAGATCCCGCTTTTCACAAAAGGGGCAACAGCTT 486
 QY 162 ValIyValAlaLeuValAlaLeuAspThrIySileuAsnIySileuValGluIleuArgGlyIle 181
 DB 487 GTTAAAGCAGTGGCGCTTGATACAGACGAAATAAAGAGTGCATTAAGAGGCATC 546
 QY 182 GluGluIleAlaGlnIyValAlaLeuAspAspValIleSlyTyIleThrAlaIyAspProGlu 201
 DB 547 GAACAAATCGCACAAATTCGCAATTAAGCAATGATGTCTATATATTAACGCCAAGCCTGAG 606
 QY 202 TyIyIyValIleMetAspAspValAlaArgGlyIleValIyValAlaAspValAlaGlnSerSer 221
 DB 607 TATTAAGGTGAATGAATGATGTGCGGTGAATTTGCAAGCGGATGTGCTCAGAGCAGC 666
 QY 222 TyrGlyLeuTyGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
 DB 667 TACGGTGTGATGACAAAGACAGATCTTAGCGGTGCGCATACAGGGCTTGATACAGGT 726
 QY 242 ArgAsnAspSerSerMetIyGluAlaPheArgGlyIySileuThrAlaLeuTyAlaLeu 261
 DB 727 CGCAATGACAGTTCATGATCAGTAGAGCTTCGCGGGGAAAATTAATGCAATTATATCATTTG 786
 QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyIleSlyTyIleThrIleValAlaGlySerVal 281
 DB 787 GACCGAGCAAGAAATTAAGCAATGATACGAATGTCTATGATGCGACATGTGGCTCCGCTA 846
 QY 282 LeuGlyAsnGlyAlaThrAsnIyGlyMetAlaProGlnAlaAsnIyValPheGlnSer 301
 DB 847 TTAGGAACCGGCTCACTAAATTAAGATGCGCTCAGCGGAATCTAGTCTCCAACTT 906
 QY 302 IleMetAspSerSerGlyIyLeuGlyIyLeuProSerAsnLeuGlnThrIleuPheSer 321
 DB 907 ATCATGATAGTACGGGTGGGACTTGGAGAGACTTACCTTGAAATCTGCAAACCTTATTCAGC 966
 QY 322 GlnAlaPheSerAlaGlyAlaArgIleHieThrAsnSerTPGlyAlaAlaValAsnGly 341

Db 967 CAGACATACAGTGTGTCGACAGAAATTCATACAACTCCTGGGAGCAGACGATGATGGG 1026
 QY 342 AATATTTTThrAspSerArgSerValAspAspTyrValArgValAspAspMetThrIle 361
 Db 1027 GCTTACACACAGATTCACGAAATGTGAGCTATGTGCGCAAAAATGATATGACGATC 1086
 QY 362 LeuPheAlaIleValAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
 Db 1087 CTTTCGCTGCGGGAGATGAGACCGGACCAATCATGTCGACACGACAGCT 1146
 QY 382 LyAsnAlaIleThrValGlyValaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
 Db 1147 AAAAATGCATTAACAGTCGAGCTACGGAACCTCCGCGCAAGCTTTGGGCTTATGCG 1206
 QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrIleAspGlyTyrGlyIle 421
 Db 1207 GACCAATATCAACCAATGTGGCACAAGTCTCTTCACGTGACCGACAAAGGATGAGCGATC 1266
 QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
 Db 1267 AAACCGATGTCAATGGCACCAGGAACTTCATCATCATGCAAGATCTTCTTTCACCG 1326
 QY 442 AspSerSerPheTyrAlaAsnHisAspSerIlyTyrAlaTyrMetGlyGlyThrSerMet 461
 Db 1327 GATTCCTCCTCTCGGGCGCAACATGACAGTAAATATGACATACATGCGTGGAACTCCATG 1386
 QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuAspGluHisPheValIleAsnArg 481
 Db 1387 GCTACACGATGCTGTGCGAAACGTGGACAGCTTCGAGACATTTGTGAAAACAGA 1446
 QY 482 GlyIleThrProLysProSerLeuLeuValAlaLeuIleAlaGlyValaAlaAspVal 501
 Db 1447 GGCATCACACCAAAAGCTTCTATTAATAAGGGGACATGATGGCGGGGACGTACATC 1506
 QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyValaThrLeuAspLysSerLeu 521
 Db 1507 GGCCTTGGTACCCCAACGCTAACAGAGATGGGACGATGACATTTGATTAATCCCTG 1566
 QY 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnValaThrTyrThr 541
 Db 1567 AACGTGCTTATGTAAACGATCCAGTTCTCTATCCACCAAGCAAAAGCGATACCTCG 1626
 QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAla 561
 Db 1627 TTTATCGCTACGCGGCAAGCTTTGAATACTCCCTGATGATGATGCCCTCG 1686
 QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
 Db 1687 AGCACAACTGCTTCCGTAAAGCTTGTCAATGATCTGGACCTTGTCAATTAACCGCTCAAT 1746
 QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArg 601
 Db 1747 GGCACACGATGATGAAATGATCTTACTTCCATTCATGATGATTAACCTGGATGGCGCG 1806
 QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
 Db 1807 AATAACGTAAGAAATGATTTATTAATGACACCAAGCGGACGCTATTAATTAATGAATTA 1866
 QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
 Db 1867 CAGGCTTATTAACGTACCGGTGGACACAGAACTTCTCGTGGCAATGTGMAAT 1920
 RESULT 3
 AA337278
 ID AA337278 standard; DNA; 1923 BP.
 AC AA337278;
 DT 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 DE Bacillus alkaline protease encoding DNA.

KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidizing agent; 88.
 OS Bacillus sp.
 PN W0918218-A1.
 XX 15-APR-1999.
 PD 07-OCT-1998; 98WO-JP004528.
 PF 07-OCT-1998; 98WO-JP004528.
 XX 07-OCT-1997; 97JP-00274570.
 PR 07-OCT-1997; 97JP-00274570.
 XX (KAO) KAO CORP.
 PA
 XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 PI Shikata S, Nomura M;
 DR MPI, 1999-287736/27.
 DR P-P8DB; AAY17088, AAY17090.
 PT Alkali protease from Bacillus used in washing powders.
 PS Disclosure; Page 58-63; 71pp; Japanese.
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidizing agents. The
 CC alkali protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidizing agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease encoding
 CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
 XX
 SQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,35e-227 Length: 1923
 Score: 3065.00 Matches: 595
 Percent Similarity: 96.9% Conservative: 23
 Best Local Similarity: 93.3% Mismatches: 20
 Query Match: 93.8% Indels: 0
 DB: 2 Gaps: 0
 US-10-784-870-4 (1-639) x AA337278 (1-1923)
 QY 2 ArgLysIleValPheLeuSerValLeuSerAlaAlaAlaIleLeuSerThrValAla 21
 Db 7 AAGAAAGAAAAGGTTTATCTGTTTATCACTCAGCGAATTTGTGCATGTTGCG 66
 QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
 Db 67 TTAAGTATTCATCTGACGGGTGCAAGGAATTTATGATCTGGAATTCAAAGAAATTCAG 126
 QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPheLeu 61
 Db 127 ACAACACTGATGCTAAAGGTTTCTCAAGACAGGGGAGACTGCTGCTGCTTTCTG 186
 QY 62 ValGluSerGluAsnValIleValLeuLeuLysGlyLeuLeuLysValLeuGluThrValPro 81
 Db 187 GTGGAATCTGAAATGTGAAACTCCCAAAAGTTTGCAAGAAAGCTTGAACAGTCCCG 246
 QY 82 AlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGluGluThrLysGln 101
 Db 247 GCAATATATAAATCCATATTATCAATTCATATGACCAATTTTGAAGAAACAAACAG 306
 QY 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121

```

Db      307  CAGCTGGAAGAAACAGGGGAAAGATTCTGACTACACTGATTATGCTTACATTGTC 366
      122  GUYTYGUGLYAaPValGInSerLyVAlArgSerLLeGluHISValGInSerValGlu 141
      367  GAGTAAAGAGGGCGATCTTAAGTCAGCAACAGCAAGCAATTGAGCAGCTGGATCCGGAG 426
      142  ProTYLeuProLYeTYrLYsIIeAaPProGInLeuPheThrLYeGILysEThrLeu 161
      427  CCTATTTCCTCAATACAGAAATAGATCCCAAGCTTTTCACAAAAGGGGCGATCAGAGCTT 486
      162  ValLYsAlaLeuAlaLeuAaPThrLYsGInAsnAsnLYsGluValGInLeuArgGILys 181
      487  GTAAGAGAGTGGCGCTTGATACAAAGCAAGAAATTAAGAGGTGCATTTAAGAGGCATC 546
      182  GUGLUUILLaGInTYrValAlaSerAaAaPValHISTYrLLeThrAlaLYsProGlu 201
      547  GAACAATCGCAATTCGCAATAGCAATGATGTCTATATATATATATATATATATATATAT 606
      202  TYrLYsValMeTAaAaPValAlaArgGILysValLYsAlaAaPValAlaGInSerSer 221
      607  TATTAAGTATAGATAGATGTGCGCGTGAATGTCAAAGCGGATGTGCTCAGAGCAGC 666
      222  TYrGILysTYrGILysGInLYsGInLYsAlaValAlaAaPThrGILysLeuAaPThrGILys 241
      667  TACGGGTTGATAGACAGAGCAGATCGTAGCGGTTCCGATACAGGCTTGATACAGGT 726
      242  ArgAaAaPSerSerMeThISGLysAlaPheArgGILysLYsLLeThrAlaLeuTYrAlaLeu 261
      727  CCCAATGACAGTTCGATGATGAGAACCTTCGCGGGAATAATTACTCATTAATATGATG 786
      262  GLYArgThrAaAaPValAaAaPThrAaGILysGILysThrHISValAlaGILysSerVal 281
      787  GAGCGAGCAATTAATGCAATGATGACGAATGTCAGTAGTCGATGTGCTGCTCCGTA 846
      282  LeuGILysGILysAlaThrAaAaPValMeTAaAaPProGILysAlaAaAaPValPheGInSer 301
      847  TTAGGAAAGCGCTCCCTAATTAAGGAATGGCGCTCAGCGCAATTAAGCTTTCCAAATC 906
      302  ILeuAaPSerSerSerGILysLYsGILysLYsLeuProSerAaAaPInThrLeuPheSer 321
      907  ATCATGATAGCGGTGGGAGACTTGGAGACTTCCATCTCGAATCTGCAAACTTAATTCAGC 966
      322  GlnAlaPheSerAlaGILysAlaArgGILysThrAaAaPThrGILysAlaValAlaGILys 341
      967  CAGCATACAGTGTGCTGCTCCAGAAATTCATCAAACTCTCGGAGACACAGAGATGG 1026
      342  AlaTYrThrThrAaPSerArgAaAaPValAaAaPValArgLYsAaAaPMeThrIle 361
      1027  GCTTACCAACAGATTCAGAAATGTGATGACTATGTGCGCAAAATGATATGACGATC 1086
      362  LeuPheAlaAlaGILysGILysArgProAaGILysLYsThrIleSerAlaProGILysThrAla 381
      1087  CTTTTCGCTGCGGGAATGAAGGACCGAAGCGGGAACCATCATGTCACACAGCAGCT 1146
      382  LYsAaAaPValIleThrValGILysAlaThrGILysLeuAaAaPProSerPheGILysTYrAla 401
      1147  AAAAAGCATATACAGTGTGAGCTAGGAAACCTTCGCGCAAGCTTGGGCTTATGGG 1206
      402  AaAaPValIleAaAaPValAlaGInPheSerSerArgGILysProThrLYsAaPValArgIle 421
      1207  GACATATACCAATGTGGCAGACTTCTTTCACGTGGACCGACAAAGATGAGCGGATC 1266
      422  LYsProAaPValMeTAaAaPProGILysTYrIleLeuSerAlaArgSerSerLeuAlaPro 441
      1267  AAAACCGATGTATGCGACCGGGAAGCTTCAATATACGAAAGATCTTCTGTGACCG 1326
      442  AaPSerSerPheThrAlaAaAaPValAaAaPValArgLYsTYrAlaTYrMeGILysLYsThrSerMeT 461
      1327  GATTCCTCTTCTGGGCGAAGCATGACAGTAATATATGATACATGAGTGAAGCTCCAGT 1386
      462  AlaThrProIleValAlaGILysAaAaPValAlaGInLeuArgGILysIlePheValLYsAaAaP 481

```

```

Db      1387  GCTACACGATCGTGTGCGTGAAGACGAGCAGCTTCGTGAGCATTTGTGAAAAACAGA 1446
      482  GILysIleThrProLYsProSerIleuLYsAlaAlaLeuIleAlaGILysAlaAaPVal 501
      1447  GGCATCACACCAAGCTTCTTATTAAGGGGACGTATGCGCGTACCTGACATC 1506
      502  GILysGILysTYrProAaAaPValAaAaPValTYrArgValThrLeuAaPValysSerLeu 521
      1507  GGCCTTGCTTACCGCAACGATACCAAGATGGAGCGATGATCATTTGATTAATATCTCTG 1566
      522  AaAaPValAlaTYrValAaAaPValysSerAlaLeuSerThrSerGILysAlaThrTYrThr 541
      1567  AACGTTGCTTATGTAAGAGTCCAGTCTCTATCCACAGCCAAAGCGATCTCTCG 1626
      542  PheThrAlaThrAlaGILysLYsProLeuLYsIleSerLeuValITPserAaAaPProAla 561
      1627  TTTACTGCTTACGCGGCAAGCTTGAATACTCCGTGTGTGTATGTATGCTATGCTCCCTGG 1686
      562  SerThrThrAlaSerValThrLeuValAaAaPValAaAaPValIleThrAlaProAa 581
      1687  AGCAACACTGCTTCCGTAAAGCTTGTCAATGATCTGAGCTTGTCAATTCAGCTCCCAAT 1746
      582  GILysArgTYrValGILysAaAaPSerAlaProPheAaAaPValAaAaPValArg 601
      1747  GGCACACAGTATGTAAGAAATGATCTTCTGCAATCAATGAATGATGATGATGATGATG 1806
      602  AaAaPValGILysAaAaPValPheIleAaAaPProGInSerGILysThrIleGILysVal 621
      1807  AATTAAGTAAAGAAATGATTTATTAATGACACCAAGGCGGACGTATCAATGATGATG 1866
      622  GlnAlaTYrAaAaPValProValGILysProGInAaAaPProSerLeuAlaIleValAa 639
      1867  CAGGCTTATACGATCCGCTGACCAAGACCTTCTGCTGCAATGTGAAAT 1920

RESULT 4
ADM40772
ID      ADM40772 standard; DNA; 1923 BP.
XX
XX      ADM40772;
XX
XX      01-JUL-2004 (first entry)
XX
XX      DNA encoding alkaline protease from Bacillus sp. KSM-KP43.
XX
XX      alkaline protease; laundry detergent; bleaching agent; detergent;
XX      denature-cleaning agent; ds; gene.
XX
XX      Bacillus sp.; KSM-KP43.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..1923
XX      FT      /*tag= b
XX      FT      /product= "Bacillus sp. KSM-KP43 alkaline protease"
XX      FT      sig_peptide 1..618
XX      FT      /*tag= a
XX      FT      /product= "Alkaline protease prepro sequence"
XX      FT      mat_peptide 619..1923
XX      FT      /*tag= c
XX      FT      /product= "Mature alkaline protease"
XX
XX      US2004072321-A1.
XX
XX      15-APR-2004.
XX
XX      09-JUN-2003; 2003US-00456479.
XX
XX      PF      26-JUN-2002; 2002JP-00186387.
XX
XX      PR      18-OCT-2002; 2002JP-00304232.
XX
XX      (KAOS ) KAO CORP.
XX
XX      Sato T, Okuda M, Takimura Y, Sumitomo N, Nomura M, Kobayashi T;
XX

```

DR MPI; 2004-328572/30.
DR P-PSDB; ADM40773, ADM40770, ADM40771.
XX New alkaline protease having a mutant prepro sequence where amino acid
PT residues at positions 52, 75 and 142 are substituted with another amino
PR acid residue, useful as enzyme component of laundry detergents, or
XX bleaching agents.
XX
PS Example 1; SEQ ID NO 3; 29pp; English.
XX
CC The invention relates to an alkaline protease having a prepro sequence.
CC The prepro sequence is a mutant sequence of SEQ ID NO: 1, or an amino
CC acid sequence having 80% homology or higher to SEQ ID NO: 1, where amino
CC acid residues at: (a) position 52 is substituted by aspartic acid or
CC arginine; (b) position 75 is substituted by alanine or arginine; and (c)
CC position 142 is substituted by lysine. The alkaline protease is useful as
CC an enzyme which can be incorporated into laundry detergents, bleaching
CC agents, detergents for cleaning hard surfaces or drainpipes, denture-
CC cleaning agents, and detergents for sterilising medical apparatus. The
CC present sequence represents DNA encoding alkaline protease from *Bacillus*
CC sp. KSM-KP43.
XX
SQ Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,356-227 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 12 Gaps: 0
US-10-784-870-4 (1-639) x ADM40772 (1-1923)
QY 2 ArglyLVLYLVValPheLeuSerValLeuSerAlaAlaAlaLeuSerThrValAla 21
DB 7 AAGAGAAAGAGGTGTTTATCTGTTTATCAGCTGAGCGATTGTCGACTGTGCG 66
QY 22 LeuAenAenProSerAlaGlyValAlaArgThrPheAenLeuAenPheLygLyLeGln 41
DB 67 TTAAGTAATCCATCGCAGGTGGTCAGAGAAATTTGATCTGGATTTCAGAGAAATTCAG 126
QY 42 ThrThrThrAenValSerGlyPheSerLygGlnArgGlnThrGlyAlaAlaAlaPheLeu 61
DB 127 ACAACATCGATGCTAAAGGTTTCTCCAGCAGGGGCGACATGAGTGGCTGCTTTTCTG 186
QY 62 ValGluSerGluAenValLyLeuLeuLygLyLeuLeuLygLyValGluThrValPro 81
DB 187 GTGGAATCTGAAGAAATGTGAACCTCCAAAGGTTTGCAAGAGAGCTTGAACAGTCCG 246
QY 82 AlaAenAenLyLeuHleGlnValGlnPheAenGlyProGlyLeuGlnGlnThrLygGln 101
DB 247 GCAAATTAATTAATCTCAATATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 306
QY 102 LySLeuGlnThrThrGlyAlaLySLeuAenLyLeuAenLyLeuProAenLyAlaThrLyLeuAl 121
DB 307 CAGCTGGAAGAAAACAGGGGCAAGATTTCTCCAGCTCAATTCGATTATGCTTACATTGTC 366
QY 122 GlnTyrgGlnGlyAenValGlnSerLySValArgSerLyGlnHleValGlnSerValGln 141
DB 367 GAGGTATGAGGGGCGATGTTAAGTCAGCAACAGACCATTTGAGCAGTGGAAATCCGTGAG 426
QY 142 ProTyrgLeuProLySValLySLeuAenProGlnLeuPheThrLygLyAlaSerThrLeu 161
DB 427 CCTATTGTCGATTAACAGATAGATCCCGAGCTTTTCACAAAAGGGGCGATCAGAGCTT 486
QY 162 ValLySAlaLeuAlaLeuAenProThrLySLeuAenLySValGlnLeuAenLySLeu 181
DB 487 GTAAAGAGAGTGGCGCTTGATCAAGAGCAAGAAATTAAGAGTGCATTTAAGAGGCAATC 546
QY 182 GlnGlnLyLeuAlaGlnTyrgValAlaSerAenAenValHleTyrgLeuThrAlaLySProGln 201
DB 547 GAACAAATTCGACAAATTCGCAATATAGCAATATGATGCTATATTAATTAAGGCAAGCTGAG 606

QY 202 TyrLySValMetAenAenValAlaArgLyLyLeValLySAlaAenValAlaGlnSerSer 221
DB 607 TATAGGTGAAGTAATGATGTTGGCGGTGAATTTGTCAAGGGAGTGGCTCAAGAGCAGC 666
QY 222 TyrGlyLeuTyrgLyGlnGlyGlnLyLeValAlaValAlaAenThrGlyLeuAenThrGly 241
DB 667 TACGGGTTGTATGACAAAGAGCAGATCGTAGCGGTTCCGATACAGGGCTTGATACAGCT 726
QY 242 ArgAenAenSerSerMetHleGlnAlaPheArgLyLySLeuThrAlaLeuTyrgAlaLeu 261
DB 727 CGCAATGACAGTGCATGCATGATGAAGCCTTCGCGGAAATTAATTCGATTATATGCAATTG 786
QY 262 GlnTyrgThrAenAenAlaAenAenThrAenGlyHleGlyThrHleSValAlaGlySerVal 281
DB 787 GAGAGGAGCAATTAATGCAAGATGCAATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuGlyAenGlyAlaThrAenLySLeuMetAlaProGlnAlaAenLeuValPheGlnSer 301
DB 847 TTAGGAACGGCTCCATTAATTAAGAAATGGCGCTTCAGGGCAATTAATGCTTCAATCT 906
QY 302 IleMetAenSerSerGlyLyGlyLeuGlyLyLeuProSerAenLeuGlnThrLeuPheSer 321
DB 907 ATCATGATAGCGGTGGGGGACTTGAGAGACTTCTGCAATTCGAAACCTTATTCAGC 966
QY 322 GlnAlaPheSerAlaGlyAlaArgLyLeHleThrAenSerThrGlyAlaAlaValAenGly 341
DB 967 CAACGATCAAGTGTGTGGCCAGAAATTCATCAAACTCTGGGAGAGCAGCTAAATGGG 1026
QY 342 AlaTyrgThrThrAenSerAenValAenAenTyrgValArgLySAlaAenMetThrLe 361
DB 1027 GCTTACACACAGATTCAGAAATGTGATGATCATGTGGCCAAATTAATGATATGACATC 1086
QY 362 LeuPheAlaAlaGlyAenGlyAenGlyProAenGlyLyGlyThrIleSerAlaProGlyThrAla 381
DB 1087 CTTTTCCTGCGGGAATGAAGACGCAACGCGGAAACATCAATGACGACGACAGCT 1146
QY 382 LyAenAlaIleThrValGlyAlaThrGlnAenLeuAenProSerPheGlySerThrAla 401
DB 1147 AAAAATGCAATTAACAGTGGAGCTTACGAAACCTCCGCAACCTTGGGCTTATAGCG 1206
QY 402 AspAenIleAenHleSValAlaGlnPheSerSerArgLyProThrLySArgLyArgIle 421
DB 1207 GACAAATTAACATGAGGACAGTCTCTTCAAGTGGACCGAAGAGATGAGCGGATC 1266
QY 422 LySProAenValMetAlaProGlyThrTyrgLeuSerAlaArgSerSerLeuAlaPro 441
DB 1267 AAACCGAGTGCATGAGCACCGGGAACGTTCACTATCAGCAAGATCTTCTTGCACCG 1326
QY 442 AspSerSerPheThrAlaAenHleAenSerLySValTyrgMetGlyGlyThrSerMet 461
DB 1327 GATTCCTCTCTGAGGCAACCATGACAGATTAATCATCATCATGAGGATGAGGATCCCATG 1386
QY 462 AlaThrProIleValAlaGlyAenValAlaGlnLeuLyrgLySLeuPheValLySAlaArg 481
DB 1387 GCTACACGATGCTGCTGGAAGAGTGGCACAGTGTGTGAGCATTTTGGAAAAACAGA 1446
QY 482 GlyIleThrProLySProSerLeuLySValAlaLeuIleAlaGlyAlaAlaAenVal 501
DB 1447 GGCATCACACCAAGCCTTCTCTTAAGCGGCACTGATGTCGGGTGCGAGCTGACATC 1506
QY 502 GlyLeuGlyTyrgProAenGlyAenGlnGlyThrGlyTyrgValThrLeuAenLySLeu 521
DB 1507 GGCCTTGCTTACCGGAACGGAACCAAGAGTGGGAGAGTGAATGGAAATTAATCCCTG 1566
QY 522 AenValAlaTyrgValAenGlySerSerAlaLeuSerThrSerGlnLySAlaThrTyrgThr 541
DB 1567 AACCTTGCTTATGAGAGAGTGCAGTCTCTATCCACAGCCAAAGAGCACTACTCG 1626
QY 542 PheThrAlaThrAlaGlyLySProLeuLySLeuSerLeuValTyrgSerAenAlaProAla 561
DB 1627 TTTACTGCTACTGCGGCAAGCCTTTGAAATCTCCCTGATGTGTGATGCCCCCTGCG 1686

QY 562 SerThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
DB 1687 AGACAACTGCTCCGTAAACGCTTGTCAATGATCTGACCTTCTCATACCGCTCCAAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAntTpaAspGlyArg 601
DB 1747 GGCACACAGATATGTAGAAATGACTTACTTCCCATACATATGATTAATCGAGATGCGCCG 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
DB 1807 AATAACGTAGAAATATATTTATTAATGACCAACAAAGCGGACGATACATTTGAGGTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1867 CAGGCTTAATACGTACCGGTGACACAGACCTTCTCGTTGGCAATGTGAAT 1920
RESULT 5
ADSI4428
ID ADSI4428 standard; DNA; 1923 BP.
XX
XX ADSI4428;
AC
XX 30-DEC-2004 (first entry)
XX
XX Bacillus alkaline protease KP43 gene SEQ ID NO:2.
XX
XX protease; enzyme; alkaline protease; laundry detergent; KP43; ds; gene.
XX
XX Bacillus sp.; KSM-KP43.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..1923
XX FT /tag= B
XX FT sig_peptide 1..618
XX FT /product= "alkaline protease"
XX FT mat_peptide 619..1920
XX FT /tag= C
XX
XX EPI466962-A1.
XX
XX 13-OCT-2004.
XX
XX 08-APR-2004; 2004EP-00008604.
XX
XX 10-APR-2003; 2003JP-00106709.
XX
XX (KAOS) KAO CORP.
XX
XX Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;
XX
XX WPI; 2004-711313/70.
XX
XX P-PSDB; ADSI4437, ADSI4427.
XX
XX New engineered alkaline protease, useful particularly in laundry
XX PT detergents, comprising specified amino acids at particular positions.
XX
XX PS Disclosure; SEQ ID NO 2; 31pp; English.
XX
XX The invention relates to a novel alkaline protease. The new alkaline
XX protease comprises an amino acid sequence in which one or more amino acid
XX residues selected from those located at 7 specific positions within
XX CC ADSI4427, or at positions corresponding to these positions are: position
XX 15 (thiolidine), position 16 (threonine or glutamine), position 166
XX CC (glycine), position 167 (valine), position 187 (serine), position 346
XX CC (arginine), and position 405 (aspartic acid). The alkaline protease is
XX useful in industry particularly in laundry detergents, but also e.g. in
XX fibre modifying agents, leather processing agents, cosmetic compositions,
XX CC bath additives, food-modifying agents, and pharmaceuticals. The present
XX CC sequence encodes the wild-type Bacillus sp. KSM-KP43 alkaline protease,
XX KP43.
XX
XX Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,356-227 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 13 Gaps: 0
US-10-784-870-4 (1-639) x ADSI4428 (1-1923)
QY 2 ArgLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
DB 7 AAGAGAAAAGAGTGTATTTATCTGTTTATACGCTGCGACGATTTTGTGACTGTGG 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
DB 67 TTAAGTAAATCAATCTGACAGTGTGCAAGAAATTTGATCTGGAATTCAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaPheLeu 61
DB 127 ACAACAACAGATGCTAAAGGTTTCTCAAGCAGGCGCAGACTGTGCTGCTTTCG 186
QY 62 ValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysGlyLysThrValPro 81
DB 187 GTGAAATCTGAATAATGTGAACCTCCAAAGGTTTGCAAGAAAGCTTGAACAGTCCG 246
QY 82 AlaAsnAsnLysLeuHilIleValGlnPheAsnGlyProIleLeuGluGluThrLysGln 101
DB 247 GCATAATTAATTAATCTCAATATTAATCAATTCATATTCATATTCATATTCATATTCAT 306
QY 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
DB 307 CAGCTGAAATAAAGCGGCGCAAGATTCGACTCATACATTCGATTAATTCATATTCG 366
QY 122 GluTyrGluGlyAspValGlnSerLysValArgSerIleGluHilValGluSerValGlu 141
DB 367 GAGTATGAGGCGGAGTTTATGATGACGACACACACATTCATGACAGTGAATCCGTGG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThrLeu 161
DB 427 CCTTATTTGCGGATTAATGATTAATGATCCCGCTTTTCACAAAAGGCGCATCAGGCTT 486
QY 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
DB 487 GTAAAGACAGTCCCTGATCAAAAGCAAAATAAAGAGTCAATTAAGAGGCATC 546
QY 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValHilTyrIleThrAlaLysProGlu 201
DB 547 GAACAATGCAACAATTCGCAATTAAGCAATGATGTGCTTAATTAACGCAAAAGCCTGAG 606
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
DB 607 TATAAGGTGATGAATGATCTTCCGCTGGAATTCGAAAGCGGAGTGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
DB 667 TACGGGTGTATGACAAAGACAGATCGATCGGCTTGCCGATACAGGCTGATCAGGT 726
QY 242 ArgAsnAspSerSerMetHilGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
DB 727 GCAATATGACAGTTCATGATGATGACCTTCGCGGAAATTAATGCAATTAATGATCATTTG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrLysGlnGlyThrHilValAlaGlySerVal 281
DB 787 GAGACGAGAAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
DB 847 TTAGGAAACGCTCCATTAATTAAGAAATGCGGCTCAGGCGCAATCTGATCTCAATCT 906
QY 302 IleMetAspSerSerGlyGlyLeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSer 321

OY		62	VAlGluSerGIuAsnValIysLeuLeuLysGIyLeuLeuLysLeuGIuThrValIPro	81
Db		187	GtGGATCTCGAAAGTGTGAACCTCCCAAAAAGTTTGCGAAMAGAAGCTTGAAACAGTCCCG	246
OY		82	AlaAsnAsnLysLeuHisIleValGlnPheAsnGIyProIleLeuGlnGluThrLysGln	101
Db		247	GCAAAATTAAATCCTCATATTATTCCAATTCAATGACAACAATTTTAGMAAAAAACAACAG	306
OY		102	LysLeuGIuThr-ThrGIyAlaLysIleLeuAspTYrIleProAspTYrAlaTYrIleVal	121
Db		307	CAGCTGGAAAAAACGGGGCCAAGANTCTCGACTACATACCCTGATTATTAAGCTTACATGTGC	366
OY		122	GIuTYrGIuGIyAspValGlnSerLysValArgSerIleGluHisValGIuSerValGIu	141
Db		367	GAGTATGAGGGCGGAGTGTAACTTAACTCACACACACACACATTTGACACGTGGAACTCCGTGAG	426
OY		142	ProTYrLeuProLysTYrLysIleAspProGlnLeuPheThrLysGIyAlaSerThrLeu	161
Db		427	CCTTAATTTGCCGATATACAGAAATGAATCCCCAGCTTTTCACAAAAGGGGCATCAAGCTT	486
OY		162	ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGIuLeuTYrGIyIle	181
Db		487	GTAAGAAGCAGTGGCCCTTGATACAAAGCAGAAAAATAAGAGGTGCATTATTAAGAGCATTC	546
OY		182	GIuGIuIleAlaGIuTYrValAlaSerAsnAspValHisTYrIleThrAlaLysProGlu	201
Db		547	GAAACAATTCGCAACAATTGCGCAATTAAGCATGATGTGCTATATTATTCGCAAAAGGCTGAG	606
OY		202	TYrLysValMetAsnAspValAlaArgGIyIleValLysAlaAspValAlaGlnSerSer	221
Db		607	TATAAGGTGAATGAATGATGTGGCGGTGAATTTCAAAGCGAGTGGCTCAGACAGAC	666
OY		222	TYrGIyLeuTYrGIuGIuGIyGlnIleValAlaValAlaAspThrGIyLeuAspThrGIy	241
Db		667	TACGGATTGTATGACAAAGACAGATCGTAGGGTGTCCGATACAGGGCTGTGATACAGT	726
OY		242	ArgAsnAspSerSerMetHisGIuAlaPheArgGIyLysIleThrAlaLeuTYrAlaLeu	261
Db		727	CGCAATGACAGTTGCATGACATGAACCTTCCGCGGAAATTAATCTGCATTATATGACATTC	786
OY		262	GIyArgThrAsnAsnAlaAsnAspThrAsnGIyHisGIyThrHisValAlaGlySerVal	281
Db		787	GGACCGAAGCAATTAATGCCAATGATTCGAATGGTCATGGTAGCAGATGGGTGGCTCCGTA	846
OY		282	LeuGIyAsnGIyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer	301
Db		847	TTAGGAAACGGCTCCCATTAATAAGAAATGGCGCTCAGGGGCAACTAGTCTTCCAACTCT	906
OY		302	IleMetAspSerSerGIyGIyLeuGIyGIyLeuProSerAsnLeuGlnThrLeuPheSer	321
Db		907	ATCATGATAGTACGGGTGGGGGACTTGAAGACTACCTTGAAATCTGCACAACTTATTACGC	966
OY		322	GlnAlaPheSerAlaGIyAlaArgIleHisArgThrAsnSerTPGIyAlaAlaValAsnGIy	341
Db		967	CAAGCATACAGNGCTGGTGCACGAATTCATCAAACTCTCGGGAGCAGCAGTGAATGG	1026
OY		342	AlaTYrThrTrpAspSerArgAsnValAspAspTYrValAlaGlyAsnAspMetThrIle	361
Db		1027	GCTTACACAAACGATTTCCAGAAATGTGATGACTATGTGCCCAAAAGATATGACGATTC	1086
OY		362	LeuPheAlaAlaGIyAsnGIyLysProAsnGIyGIyThrIleSerAlaProGIyThrAla	381
Db		1087	CTTTTCGTCGCGGAAATGAAGACCCGAAACGGCGGAAACATCAGTGACACAGGCAACACT	1146
OY		382	LysAsnAlaIleThrValGIyAlaThrGluAsnLeuArgProSerPheGIySerTYrAla	401
Db		1147	AAAAATGCATATACAGTCGAGCTACGAAAAACCTCCGCCCAAGCTTTGGGTCTTATACG	1206
OY		402	AspAsnIleAsnHisValAlaGlnPheSerSerArgIyProThrLysAspGIyArgIle	421
Db		1207	GACAAATATCAACAGTGTGCAcAGTTCTCTTCACTGTGACCGACCAAGAGATGAGCAGATTC	1266

QY	422	lysPcRoApvna1Me1a1ProG1yThrTyr1le1uSerA1aXySere1eu1aPro	441
Db	1267	AAACCGAGATGCATGACACCGGGAAAGTTCAATACATACGAAGAATCTTCTCTGGACCG	1322
QY	442	AapSerSerPheTTrpAlaAsnH1saPSeSerLyS1yA1aTyMeG1yG1yThrSerMet	461
Db	1327	GATTCCTCCTTCTGGGCGAACCATACACTAATATGCACTACATGGGTGGAAAGTCCATG	138
QY	462	AlaThrPro1leVal1aG1yAanVal1aG1leuArgG1uH1SphVal1ySaAanArg	481
Db	1387	GCTACACCGATGTTGCTGGAAACGTGGGACAGCTTCGTGACATTTTGTGAAAAACAGA	144
QY	482	Gly1leThrProLyPProSer1eu1ySa1a1a1eu1le1aG1yAla1aAspVal	501
Db	1447	GGCATCAACCAACCAACCTTCTCTATTAAGGCGACATGATGGCCGGTGCACCTGCATC	150
QY	502	Gly1leuG1yTyProAsnG1yAanG1nG1yTTrG1yArgVal1Thc1euAspLySere1eu	521
Db	1507	GGCCTTGGCTACCCGACACGCTACCAAGATGGGAGCAAGTACATTTGATTAATCCCTG	156
QY	522	AanVal1aTyYVal1aSnG1uSerSerA1a1euSerThrSerg1nLySa1aThTyTThr	541
Db	1567	AACGTGGCTATGTAAGACAGTCCAGTCTCTATCCACCAAGCCAAAAGCAGTACTCG	162
QY	542	PheThrAlaThrAlaG1yLySPro1euLyS1leSer1euVal1TyPSetaSp1aPro1a	561
Db	1627	TTTATGCTACTAGCCGGCAGCGCTTGAAGAAATCTCCCTGGTATGCTGATCCCTCGCG	168
QY	562	SerThrThrAlaSerVal1Thc1euVal1aSnAsp1euAsp1euVal1leThrAlaProAn	581
Db	1667	AGCACACCTGCTCCGTAACGCTTGCAATATCTGACCTTGCAATTAACGCTCCAAAT	174
QY	582	GlyThrATGTYVal1G1yAanAspPheSerA1aProPheAspAanATrPaSpG1yArg	601
Db	1747	GGCACACAGTATGTAAGAAATGACTTTACTTCGCCATACATGATACATGGGATGGCCGC	180
QY	602	AanAanVal1G1uAanVal1Phe1leAanSerProG1nSerg1yThrTyTThr1leG1uVal	621
Db	1807	AATAACGTAAAGAAATGTAATTTATTAATGCAACCAACAAAGCGGAGCTAATACATTTAGGTA	186
QY	622	G1uAlaTyTrAanVal1ProVal1G1yProG1nAanPheSer1euAl1leVal1aSn	639
Db	1867	CAGGCTTATTAACGTACCGGTTGGACACACAGACCTTCTGTTGGCAATGTGTAAAT	1920
RESULT 7			
AAV82382			
ID	AAV82382 standard; DNA; 3003 BP.		
XX	AAV82382;		
XX			
DT	12-APR-1999 (first entry)		
XX			
DS	Bacillus JPI70 protease gene.		
XX			
KW	Protease; detergent; surfactant; leather processing; debittering;		
KW	flavour; ss.		
OS			
XX	Bacillus sp.		
XX			
Key	Location/Qualifiers		
FT	846..2771		
FT	/*tag= a		
FT	sig_peptide		
FT	846..944		
FT	/*tag= b		
FT	mat_peptide		
FT	1470..2768		
FT	/*tag= c		
PN	M09856927-A2.		
XX			
PD	17-DEC-1998.		
XX			
PF	09-JUN-1998; 98WO-US012005.		

PR 12-JUN-1997; 97US-00873479.
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX Sloma A, Christianson L;
XX MPI, 1999-080908/07.
DR P-PSDB; AAM89547.
PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
PS diethwashing detergents and for leather processing.
XX Claim 11; Page 52-53; 77pp; English.
XX This nucleotide sequence encodes a novel protease (see AAM82382) of
CC *Bacillus* sp. JP170 (NCIB 12513). The sequence in plasmid p170BAN is
CC contained in *Bacillus subtilis* LC20 NRRL B-21680. The protease gene was
CC isolated from chromosomal DNA of JP170 following preparation of probes
CC based on protease N-terminal and internal peptides (see AAM89549-50),
CC screening of chromosomal libraries, isolation of the 3' end of the gene
CC by inverse PCR (see AAV82410-11), reconstruction of 5' and 3' ends and
CC PCR amplification (see AAV82412-16). Claimed recombinant host cells can
CC be used in a method for producing the protease. The protease is used in
CC laundry and dishwashing detergents, for institutional and industrial
CC cleaning, and for leather processing, as well as for debittering and
CC enhancing the degree of hydrolysis of protein hydrolysates, for flavour
CC development through hydrolysis of proteins, degradation of undesired
CC peptides and in enzymatic synthesis of peptides. It has enhanced
CC stability towards oxidation under alkaline conditions, e.g. towards
CC bleaching agents of the peroxy type. The invention also provides mutant
CC cells in which the protease activity is diminished. Such cells can be
CC used for the production of heterologous recombinant proteins
XX
SQ Sequence 3003 BP; 982 A; 504 C; 645 G; 872 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,096-206 Length: 3003
Score: 2799.00 Matches: 541
Percent Similarity: 92.4% Conserva: 52
Best Local Similarity: 84.3% Mismatches: 45
Query Match: 85.7% Indels: 4
DB: 2 Gaps: 2

US-10-784-870-4 (1-639) x AAV82382 (1-3003)
QY 1 MetArgLys-----LysLysValPheLeuSerValLeuSerAlaAlaIleLeu 17
DB 846 ATGAGAAAGAAAGATCGAAGAGCGTTTATCCGTTTATCAGTTCGCACTATG 905
QY 18 SerThrValAlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
DB 906 TCTTGTGTGCTTAAAGCATCTCTTACTATTTGGGGCAACATTTGGAATGGACTTT 965
QY 38 LysGlyIleGlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAla 57
DB 966 AAGGGGATAGAGACACTTACGCTAAGAAAGCTGCCACCAAGCAAGAAAAAGGGAAG 1025
QY 58 AlaAlaPheLeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLeu 77
DB 1026 GCATCTTTCTTGTAACTCTGAAATATGTAATCCCAAGAGATTCCAAAAGAACTA 1085
QY 78 GluThrValProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlu 97
DB 1086 GAAGTAGTTCCAGCGGATTAACAAGCTATATATCGTTCAATTGAGGAGACTATTTTAAAG 1145
QY 98 GluThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyr 117
DB 1146 GAAGCGCACTTCAACTAGAGAAAGCGGAGCGAAATTCGATTACATACCAATTAAC 1205
QY 118 AlaTyrIleValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisVal 137
DB 1206 GCTTATATGTGCGAATATGATGAGGATGTAAAGCGCTAATCAAGCAATTCGCAATTG 1265

QY 138 GluSerValGluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGly 157
DB 1266 GAATCGGTGAAACCATTTATCTTTATATTAATTAACCCGCAATTTTCCAGAGCA 1325
QY 158 AlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlyValGln 177
DB 1326 GCTTCGAATTAGAGAAACAAGTAGCTTTAGATTAAGAAAGCAAGAGTAAGAGTAGCT 1385
QY 178 LeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThr 197
DB 1386 TTAAGAGATTAAGAAATATGCGCAATTAAGCCGCAATTAATATGATATATAGTACACC 1445
QY 198 AlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspVal 217
DB 1446 CCAAGCGCTGAATTCGAAGTTTGAATGACGTGCCCGGTGCATTGGAAAGCAGACGTC 1505
QY 218 AlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGly 237
DB 1506 GCAAAATAATACCTTGGCTTATATGACAAAGACAGATTGTAGCATTGCTGATACGGG 1565
QY 238 LeuAspThrGlyArgAsnAspSerSerMetHisGluValPheAspGlyLysIleThrAla 257
DB 1566 CTTGATACAGAGAAATATGACAGTTCATGATGATGACATTCGCGGTAAAGATTACCGCA 1625
QY 258 LeuTyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisVal 277
DB 1626 CTATATGACCTGGGCAAGAAATACAGCCCAATATCCAAATGACATGAAACCAATGTT 1685
QY 278 AlaGlySerValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeu 297
DB 1686 GCTGATCTGTGTAGGAAT--GTAACAAATTAAGAGATGGCAGCAGCAACCAATCTA 1742
QY 298 ValPheGlnSerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGln 317
DB 1743 GTCTTCAATCTATATTAAGATATGATGAGGCGCTGAGAGACTACCTGCTAATCTCAAA 1802
QY 318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAla 337
DB 1803 ACATTAATTCAGTCAAGCATATATAGTCTGAGCGAAGATTCATACCAATTCATGGGGGCT 1862
QY 338 AlaValAsnGlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsn 357
DB 1863 CCAGTAACCGGTGCTATACAGACAGCTCTCGAAATGTTGATGTATGAGAAATAAT 1922
QY 358 AspMetThrIleLeuPheAlaIleAlaGlyAsnGluArgProAsnGlyTyrThrIleSerAla 377
DB 1923 GATATGACGATCTTTTGGCGCGGAAATGAGGAGCAGGTACCGGTACATATGAGCA 1982
QY 378 ProGlyThrAlaLysAsnAlaIleThrValGlyValaThrGluAsnLeuArgProSerPhe 397
DB 1983 CCAGAAACAGCAAAAATGCGAATTACAGTTGGGGCAACCGAAACCTACGTCGCAAGCTTC 2042
QY 398 GlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLys 417
DB 2043 GATCTTATGGGATATATTAATTAACATGTTGCTCAATTCCTTACAGAGGCTCACTAGA 2102
QY 418 AspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSer 437
DB 2103 GATGAGAGTATTAAGCCGAGCGTCATGCAACAGGTACGATATATCTCTGCTAGATCA 2162
QY 438 SerLeuAlaProAspSerSerPheTrpAlaAsnHisAspSerLysTyrAlaTyrMetGly 457
DB 2163 TCATTAGCTCCAGATTCCTCATTTGGGCAAAACATATATATATATATATGCTTACATGGGT 2222
QY 458 GlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPhe 477
DB 2223 GGTACTCTATAGGCTACTCCAAATTTGAGCAGGTATATGTGACAAATTAAGGAGCAATTT 2282
QY 478 ValLysAsnArgGlyIleThrProLysProSerLeuLeuLysValAlaLeuIleAlaGly 497
DB 2283 GTGAAATATAGAGGGGTAACCTTAAGCTTCCCTTTAAAGCTGCTTATATGCAAGGT 2342
QY 498 AlaAlaAspValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeu 517

```
DB 2343 GCGGGGAGTGTGACTTGCTTCCAAATGCGTAACGAGTGGGAGGTAACGTTA 2402
QY 518 AaplySerLeuAenValAlaTyValAenGlySerSerAlaLeuSerThrsGlnTyS 537
DB 2403 GATTAATATCCCTTAATATGCTCATTTGTAATGAAGACGCTTTATCAACAAGTCAAAA 2462
QY 538 AlaThrTyThrPheThraLalThraLagLyLeProLeuLyLeLeSerLeuValTpsSer 557
DB 2463 GCAACATATTCGTTTACGGCTCAAGTGTGAACCCCTTAATAATATCATTGTTGGTCA 2522
QY 558 AapAlaProAlaSerThrThraLaseValThrLeuValAenAapLeuAapLeuValIle 577
DB 2523 GATGACCGACGTAAGACGACGCGCATCACTTAAGTAAATTAAGATTAAAGCTTAAGTAATC 2582
QY 578 ThrAlaProAenGlyThraTygTyValGlyAenAapPheSerAlaProPheAapAenAen 597
DB 2583 ACTGCACCAATATGAATGAACTTAATACGTCGGAATATGACTTTACAGACCGTATGATACAT 2642
QY 598 TTPAaPGlyArgAaAenValGlyAenValPheIleAenSerProGlnSerGlyThrTyTyr 617
DB 2643 TGGGATGGGAGAAACACGTGAAATATGTTTATCAATGCTCTCAAGCGGAACGTAT 2702
QY 618 ThrIleGlyValGlnAlaTyraAenValProValGlyProGlnAapPheSerLeuAlaIle 637
DB 2703 ACGAGTCGAAGTCAGGCTTACATATGACATGACATGACGTAAGTCCGCAACCTTTCTTTAGCGATT 2762
QY 638 ValAen 639
DB 2763 GTACAT 2768

RESULT 8
ADY33779
ID ADY33779 standard; DNA; 1305 BP.
XX
AC ADY33779;
XX
DT 05-MAY-2005 (first entry)
XX
DE Bacillus species alkaline protease coding sequence.
XX
KM mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
XX
OS Bacillus sp.
XX
Key Location/Qualifiers
CDS 1..1305
FT /*tag= a
FT /product= "alkaline protease"
FT /note= "no start codon given"
FT /partial
XX
BN BP1347044-A2.
XX
PD 24-SEP-2003.
XX
PF 21-MAR-2003; 2003BP-00006472.
XX
PR 22-MAR-2002; 2002JP-00081428.
PR 06-JUN-2002; 2002JP-00155987.
PR 18-OCT-2002; 2002JP-00304230.
PR 18-OCT-2002; 2002JP-00304231.
XX
PA (KAOS ) KAO CORP.
PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
PI Kobayashi T, Nomura M;
XX
DR WPI; 2003-846540/79.
DR P-PsDB; ADY33778.
XX
XX New alkaline protease having specific amino acid residue at a specific
```

```
PT position of its amino acid sequence, useful for producing detergent
PT compositions, laundry detergent, fiber modifiers, leather-treating agents
PT or pipe cleaners.
XX
PS Disclosure; SEQ ID NO 2; 31bp; English.
XX
CC The invention relates to an alkaline protease having a fully defined
CC sequence of 434 amino acids (I) given in the specification, or an amino
CC acid sequence at least 80% homologous with (I), where an amino acid residue
CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
CC selected from 16 amino acid residues. The amino acid residues at the
CC corresponding positions are selected from: position 65: proline; position
CC 101: asparagine; position 163: histidine, aspartic acid, phenylalanine,
CC lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or
CC valine; position 170: valine or leucine; position 171: alanine, glutamic
CC acid, glycine or threonine; position 273: isoleucine, glycine or
CC threonine; position 320: phenylalanine, valine, threonine, leucine,
CC isoleucine or glycine; position 359: serine, leucine, valine, isoleucine
CC or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic
CC acid, arginine or histidine. The alkaline protease is useful for the
CC production of a detergent composition, such as laundry detergent, fiber
CC modifiers, leather-treating agents, cosmetic compositions, bath
CC additives, food modifiers and pharmaceutical compositions. The alkaline
CC protease may also be used as bleaching detergent, hard surface cleansing
CC detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing
CC cleanser for medical tools. The new alkaline protease has a more potent
CC proteolytic capacity, exhibiting excellent detergency for the removal of
CC a complex stain, and has high secretion capacity. This sequence
CC corresponds to the Bacillus sp. alkaline protease coding sequence.
XX
SQ Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1..81e-159 Length: 1305
Score: 2183.00 Matches: 418
Percent Similarity: 99.34 Conservative: 13
Best Local Similarity: 96.34 Mismatches: 3
Query Match: 66.84 Indels: 0
DB: 11 Gaps: 0

US-10-784-870-4 (1-639) x ADY33779 (1-1305)
QY 206 AenAapValAlaAargGlyTleVallyAalAapValAlaGlnSerSerTyrclyLeuTyTyr 225
DB 1 AATGATGTTGGCGGTGAATGTCMAAGCGGATGTCGACAGACGCTACCGGTGTGAT 60
QY 226 GlyGlnGlyGlnIleValAlaValAlaAapThrglyLeuAapThrglyArgAenAapSer 245
DB 61 GGCACAGACAGATGTAGCGGTGCGCATACAGGAGCTTGATACAGGTGCGCATACACT 120
QY 246 SerMetHlsgIuaLpAeArgGlyLyserIleThraLaleuTyZalaLeuGlyArgThraAen 265
DB 121 TCGATGCATGAAGCCTTCGCGGGAATTTCTGCAATTATATGATTTGGAGACGACGAAAT 180
QY 266 AenAlaAenAapThraAenGlyHlsglyThrhlyeValAlaGlySerValleuGlyAenGly 285
DB 181 AATGCCAATGATACGAATGTGATGTACATGTGCGTGGCTCCGTTATGGAACCGGC 240
QY 286 AlaThraAenGlyMwerAlaProGlnAlaAenAenValPheGlnSerIleMwerAapSer 305
DB 241 TCACATTAATTAAGGAATGCGCTCAGCGCAATCTAGCTTCATTCATATGATGATAGC 300
QY 306 SerGlyGlyLeuGlyGlyLeuProSerAenAenGlnThraLeuPheSerGlnAlaPheSer 325
DB 301 GGTGGGGGACTTGGAGGACTACCTTCGAAATCTGCAAACTTATTCAGCCAAACATACAGT 360
QY 326 AlaGlyAlaArgIleHlserThraAenSerTtgclyAlaAlaValAenGlyAlaTyThrThr 345
DB 361 GCTGTGTCAGAAATTCATACAACTCCCTGGGAGAGACGATGAATGGGGCTTACCAACA 420
QY 346 AapSerArgAenValAapAapTyTValArgLyAenAapMetThrIleLeuPheAlaAla 365
DB 421 GATTCCAGAAATGTGATGATGATATGTGCGCAAAATATATGATGATGATCTTTTGGCTGCC 480
```

```
QY 366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaValAsnAlaIle 385
Db 481 GGGATGAGAGCAACCGCGGAGCACTCACTGACCAAGGACAGCTAAATCAATTA 540
QY 386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
Db 541 ACAAGTCGAGCTACGGAAGAACTCCGCCAAGCTTTGGGTTCTATGCGGACAAATATCAAC 600
QY 406 HisValAlaGlnPheSerSerArgGlyProThrIleAspGlyValGlyIleValProAspVal 425
Db 601 CATGTGGCAGATCTCTTCACTGACCGACCAAGATGAGCGATCAAAACCGATATTC 660
QY 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
Db 661 ATGGACCGGAGACCTCACTCACTACAGAGATCTTCTTTCACCGGATCTCTCTTC 720
QY 446 TrpAlaAsnHisAspSerLeuTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
Db 721 TGGGGCAACCATGACATTAATATGCAATGAGTGGAGAGCTTCATGCGTACACGATC 780
QY 466 ValAlaGlyAsnValAlaGlnLeuArgGluHisPheValIleAsnArgGlyIleThrPro 485
Db 781 GTTGCTGGAAACGTGGCAGAGCTTCGTAGCATTTTGGAAAGACAGGACATCACACCA 840
QY 486 LysProSerLeuLeuValAlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db 841 AAGCTTCTCTATTAAGGCGGACCTGATTCGGGTGACGCTGACATGCGCTTGCTAC 900
QY 506 ProAsnGlyAsnGlnGlyTyrGlyValArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACCGTAAACCAAGATGGGAGAGATGACATTTGATTAATCCCTGAACGTTGCTAT 960
QY 526 ValAsnGlySerSerAlaLeuSerThrSerGlnValAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAGTCAGTCTCTATCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 546 AlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAlaSerThrThrAla 565
Db 1021 GCCCGACAGCTTTGAAATCTCCCTGGATAGGTGATGATGCCCTGCAAGCAATCTCT 1080
QY 566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db 1081 TCCGTAACGCTTGTCAATGATCTGGAACCTTTCATTAACGCTCCAAATGGGACACAGAT 1140
QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyValArgAsnValGlu 605
Db 1141 GTAGGAATGACTTACTTCGCAATACATGATGATGAGGATGCGCGCAATTAACCTAGAA 1200
QY 606 AsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValAlaIleTyrAsn 625
Db 1201 AATGATATTATTAATGACCAACAAAGCGGAGCGTATGACATTAAGGCTTATTAAC 1260
QY 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1261 GTACCGGTTGAGCACAGACCTTCTGTTGGAAATGTGAAT 1302

RESULT 9
ADL25803
ID ADL25803 standard; DNA; 1305 BP.
AC ADL25803;
XX
XX 20-MAY-2004 (first entry)
XX
XX Bacillus alkaline protease coding sequence.
XX
XX alkaline protease; washing agent; gene; ds; enzyme.
XX
XX Bacillus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..1305
```

```
FT /*tag= a
FT /partial
FT /product= "Bacillus alkaline protease"
FT /note= "No start codon is given"

JP2004008085-A.
XX
XX
XX 15-JAN-2004.
PD
XX
XX 06-JUN-2002; 2002JP-00165950.
PF
XX
XX 06-JUN-2002; 2002JP-00165950.
PR
XX
XX (KAO ) KAO CORP.
XX
XX WPI; 2004-094297/10.
DR
XX P-PSDB; ADL25802.
XX
XX Novel mutant alkaline protease produced by substituting the amino acid
XX residues useful as washing agent.
XX
XX Disclosure; SEQ ID NO 2; 21pp; Japanese.
XX
XX The invention comprises a mutant Bacillus alkaline protease which
XX contains substitutions at positions 163, 170 and 434. The mutant alkaline
XX protease is useful as a washing agent. The present DNA sequence encodes a
XX Bacillus alkaline protease of the invention.
XX
SQ Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,81e-159 Length: 1305
Score: 2183.00 Matches: 418
Percent Similarity: 99.3% Conservative: 13
Best Local Similarity: 96.3% Mismatch: 3
Query Match: 66.8% Indels: 0
DB: 12 Gaps: 0

US-10-784-870-4 (1-639) x ADL25803 (1-1305)

QY 206 AsnAspValAlaArgGlyIleValAlaAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db 1 AATGATGTGGCGCTGGAATGTCAAGCGGATGTGCTCAAGACACTACGGGTGTAT 60
QY 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyValArgAsnAspSer 245
Db 61 GGAACAGACGATCGTAGCGGTGGCCGATACAGAGGCTTGAATACAGGTCGCAATGACAGT 120
QY 246 SerMetIleGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db 121 TCGATGCAATGAGCTTCCCGGGAATTAATCTGATTAATGCAATTTGGACGGAACAAAT 180
QY 266 AsnAlaAsnAspThrAsnGlyValIleGlyThrIleValAlaGlySerValLeuGlyAsnGly 285
Db 181 AATGCCAATGATGAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 286 AlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
Db 241 TCCACTATATAAGGAATGGCGCTCAAGCGCAATCTTCAATCTTATCATGATGATGATGAT 300
QY 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325
Db 301 GTGGGGGACCTTGGAGGACCTACCTTCATATCTGCAACCTTATTCACGCAAGCATACAGT 360
QY 326 AlaGlyAlaArgIleHisThrAsnSerTyrGlyValAlaIleValAsnGlyAlaTyrThrThr 345
Db 361 GCTGTGCCAGAAATTCATTAACAATCTCTGGGGAGACAGATGATGGGCTTACACAAACA 420
QY 346 AspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIleLeuPheAlaAla 365
Db 421 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaValAsnAlaIle 385
```

```

Db      481 GGGAAATGAAGACCGAACCAGCAATCAATGACACGACCACTTAAATGCAATA 540
Qy      386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
Db      541 ACAAGTCGAGCTACCGAATAACCTCCGCCAAGCTTGGGCTTATGCGGACAAATATCAAC 600
Qy      406 HisValAlaGlnPheSerSerArgGlyProThrIleAspGlyArgIleLeuProAspVal 425
Db      601 CATGTGGCAGAGTCTCTTCAGGTGACCGAACAAGATGACAGGATCAACACCGAGTGTCC 660
Qy      426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerIleuAlaProAspSerPhe 445
Db      661 ATGGCACCAGGAACTTCACTATACGAACATCTTCTCTTGACCGAATCTCTCTTC 720
Qy      446 TrpAlaAsnHisAspSerIleTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
Db      721 TGGGCGAACCAGTAAATATGATCATATGGGTGGAACGTCCATGGCTACACCGATC 780
Qy      466 ValAlaGlyAsnValAlaGlnLeuArgGluHisPheValIleAsnArgGlyIleThrPro 485
Db      781 GTTGTGGAAACGTGACAGCTTCTGTAGCAATTTGTGAAAAACAGAGCATCACACCA 840
Qy      486 LysProSerLeuLeuValAlaIleuIleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db      841 AAGCTTCTCTTAAACCGGCACTGATTCGGGTGCGAGCTGACATCGGCTTGGCTAC 900
Qy      506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeuAsnValAlaTyr 525
Db      901 CCGAACGGTACCAAGATGAGGAGAGATGACATTTGATTAATCCCTGAACGTTGCTTAT 960
Qy      526 ValAsnGluSerSerAlaLeuSerThrSerGlnValAlaThrTyrThrThrAlaThr 545
Db      961 GTGAACGAGTCCAGTCTCTATCCACACCAAAAGCAGATCTGTTACTGCTACT 1020
Qy      546 AlaGlyLysProLeuLysIleSerLeuValThrSerAspAlaProIleSerThrAla 565
Db      1021 GCCGCAACCTTTGAAATCTCCGTGATGCTGATGAGCCCTGCGACACACTGCT 1080
Qy      566 SerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
Db      1081 TCCGTAAACGCTGTCAATGATCTGGACCTTGTCAATACCGCTCCAAATGGACACAGTAT 1140
Qy      586 ValGlyAsnAspPheSerAlaProPheAspAsnMetTrpAspGlyArgAsnAsnValGlu 605
Db      1141 GTAGAAATAGACTTAACTTCGCCATACATGATTACTGGATGCGCGCAATTAACGTAAGA 1200
Qy      606 AsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGlnAlaTyrAsn 625
Db      1201 AATGATTTTATTATATGACCAACAAAGCGGAGCTATATCAATTGAGGTACAGGCTTATAC 1260
Qy      626 ValProValGlyProGlnAsnAspSerLeuAlaIleValAsn 639
Db      1261 GTACCGGTTGACCAAGACTTCTCTGTTGGCAATTTGTAAT 1302

```

RESULT 10

```

AA027516
ID      AA027516 standard; DNA; 1299 BP.
AC      AA027516;
DT      05-FEB-1993 (first entry)
DE      Alkali-protease Ya enzyme gene.
XX      Alkali resistance; surface active agent resistance; detergency improver;
XX      88.
OS      Bacillus sp. Y.
XX      Key Location/Qualifiers
XX      CDS 1..1299
FT      /*tag= a

```

```

XX      XX
PN      JP04197182-A.
XX      PD 16-JUL-1992.
XX      PF 28-NOV-1990; 90JP-00327110.
XX      PR 28-NOV-1990; 90JP-00327110.
XX      PA (LION ) LION CORP.
XX      DR WPI; 1992-288440/35.
XX      P-PSDB; AAR26274.
XX      PT DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX      resistance and improves detergency.
XX      PS Claim 3; Page 2; 17p; Japanese.
CC      The sequence is that of the alkali-protease Ya enzyme gene which can be
CC      used in the recombinant production of Ya enzyme. Ya enzyme is excellent
CC      in alkali resistance and surface active agent resistance and improves
CC      detergency
SQ      Sequence 1299 BP; 401 A; 246 C; 297 G; 355 T; 0 U; 0 Other;

Alignment Scores:
Sred. No.: 1,42e-144 Length: 1299
Score: 1990.50 Matches: 381
Percent Similarity: 94.0% Conservative: 27
Best Local Similarity: 87.8% Mismatch: 25
Query Match: 60.9% Indels: 1
DB: Gaps: 1

US-10-784-870-4 (1-639) x AA027516 (1-1299)
Qy      206 AsnAspValAlaArgGlyIleValIleAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db      1 AATGATGTAAGCAAGGAGATGTAAGTAAAGCTGATGTTCACAAACATTAACGATTAAT 60
Qy      226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
Db      61 GGCACAAAGTCACATGTTGCACTAGCGACGACACAGCTTATGATACAGCTGTAAACATAT 120
Qy      246 SerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db      121 TCTATGCAATGAAGCATTCGCGGGAATATCAACAGCTCTTACGCTTGAAGAACTAAT 180
Qy      266 AsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGly 285
Db      181 AATGCAAGTATCCCAATGGGCATGCAACATGTAGCAGGTTCTGTACTTGAT--- 237
Qy      286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
Db      238 GCTTAAATTAAGAAATGGCTCCGCAAGCTTACCTTAGCTTCAATCTATTAATGATAC 297
Qy      306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325
Db      298 AGCGGAGGATTAAGGTGGCTTACCATCGAATTAATAAGCTTATTAATGCAAGCTTGAAAT 357
Qy      326 AlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGlyValaTyrThr 345
Db      358 GCTGAGCAAGAAATTCATTAATCTTGGGAGCCCAAGTAATGAGGCCGTAACCTGCT 417
Qy      346 AspSerArgAsnValAspAspTyrValaArgLysAsnAspMetThrIleLeuPheAlaIle 365
Db      418 AACTGAAACAAAGTGAAGATGAGATTTGCAATATATGATACAGCTTCTTTCAGCT 477
Qy      366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLeuAsnAlaIle 385
Db      478 GGTATATGAAGGCTTAATTCAGGAACATTAATGTGCTCCAGGTAACGCAAAATGCTATTT 537
Qy      386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405

```



```

QY 255 ILeThrAlaLeuTyr---AlaLeuGlyArgThrAsnAsnAlaAsnAsnLeuThrAsnGlyHis 273
   |||
Db 541 GTGATAGGCTGGTGAACAGCGGTCACAGCGGACGCTTACAGTACGAGGAC 600
QY 274 GlyThrHisValAlaGlySerValLeuGlyAsnGlyValaThrAsnLys-----Gly 290
   |||
Db 601 GGAACCAAGTTCGGGTATCGTTCGGGAAACCGGACGCTTAACTCCAGTACATAGG 660
QY 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
   |||
Db 661 GTCCGCCCCGCGCGAAGCTCGTCGGCGTCAAGTTCTCGGTCCGACGGTGGGAGG 720
QY 309 LeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyVala 328
   |||
Db 721 GTCTCCACCACTATCGGGGTGTGACTGGGTGTCAGAACAGGACAACTACGGGATTA 780
QY 329 ArgIle-----HisThrAsnSerTrpGly 336
   |||
Db 781 AGGGTCATCAACCTCTCCCTCGGCTCTCCAGAGCTCCGAGGAAACGACTCCCTCAGT 840
QY 337 AlaAlaValAsnGlyValaTyrThrThrAspSerArgAsnValAspAspTyrValaArgLys 356
   |||
Db 841 CAGGCGGTCAACAAACGCTGGGACGCC----- 867
QY 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArgProAsnGlyGlyThrIleSer 376
   |||
Db 868 ---GGTATAGTAGTCCGCTCGCGCGCGGCAACGCGGCGCAACCTACACCGCTCGGC 924
QY 377 AlaProGlyThrAlaValAsnAlaIleThrValGlyValaThrGluAsnLeuArgProSer 396
   |||
Db 925 TCACCCGCGCGCGCGGCAAGGTCATACGCTCGGTCA----- 963
QY 397 PheGlySerTyrAlaAspAsnIleAsnIleValaIleGlnPheSerSerArgGlyProThr 416
   |||
Db 964 ---GTTACAGCAACAGCAACATCGCAGCTTCTCCAGCAGGAGGACGAGCC 1011
QY 417 LysAspGlyArgGlyLeuProAspValMetAlaProGlyThrTyrIleLeuSerAlaArg 436
   |||
Db 1012 GCGGACGGAAGGCTCAAGCGGAAAGTCGCGCGCGGTGACATCAATAGCCCGCC 1071
QY 437 SerSerLeuAlaProAspSerSerPheThrAlaAsnIleAspSerLysTyrAlaTyrMet 456
   |||
Db 1072 GGCAGC-----GGAACAGCAGTGGGACCCGATTAACACTACCTACCAAGGCC 1122
QY 457 GlyGlyThrSerMetAlaThrProIleValaIleGly---AsnValaIleGlnLeuArgGln 475
   |||
Db 1123 TCTGGAAACGACATGGCCACCCGACGTTTCGGGCGGTGGCGGCTCATCTCCAGGCC 1182
QY 476 HisPheValLysAsnArgGlyIleThrProLys-----ProSerLeuLeuLysAlaIle 493
   |||
Db 1183 CAC-----CCGACTGGACCCCGGACAAAGGTGAAAGCCGCC 1218
QY 494 LeuIleAlaGlyAla-----AlaAspValGlyLeuGlyTyr 505
   |||
Db 1219 CTCATCTAGAACCGCGACATAGTCCGCCCAAGAGATGCGGACATCCCTACGGTGGC 1278
QY 506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu-----Asn 522
   |||
Db 1279 ---GGTAGGGTGAACGCTTACAGGCCATCAAGTACAGCAGC 1317
QY 523 ValAlaTyrValAsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThrPhe 542
   |||
Db 1318 TAGCGCAAGCTCACTTACCGGCTCCGTCGCGCAACAGGAAAGCCCGCACACTTC 1377
QY 543 ThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSerAspAlaProAlaSer 562
   |||
Db 1378 GACGCTCAGGGCGGCACTTCTGACCGGCACTTACTTGGGAC----- 1422
QY 563 ThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGly 582
   |||
Db 1423 ---ACGGGCTCGAGGACATGACCTTACTCTTACAGACCCCAAGCGG 1467

```

```

QY 583 ThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArgAsn 602
   |||
Db 1468 AACGAG-----GTTGACTACTCTTACACCGGCTTACTAC 1500
QY 603 AsnValGlyAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGln 622
   |||
Db 1501 GGCTTCGAAAGAGTCCGCTACTACAAACCGGACCGCGAACTCGACGCTCAAGTCTGTC 1560
QY 623 AlaTyrAsnValProValGlyProGlnAspPheSerLeuAlaIleValAsn 639
   |||
Db 1561 ACTACAAAG-----GGCGCGGACACTACAGTCTGACGTCGTACG 1602

RESULT 12
AAK05926
ID AAK05926 standard; DNA; 1977 BP.
XX
AC AAK05926;
XX
DT 06-MAY-1999 (first entry)
XX
DE WO9856926 Seq ID 11.
XX
KM Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
XX additive; drug; washing agent; foodstuff; chemical synthesis; de.
XX
OS Thermococcus celser.
XX
PN WO9856926-A1.
XX
PD 17-DEC-1998.
XX
PF 04-JUN-1998; 98WO-JP002465.
XX
PR 10-JUN-1997; 97JP-00151969.
XX
PA (TAKI ) TAKARA SHUZO CO LTD.
XX
PI Takakura H, Morishita M, Shimoto T, Asada K, Kato I;
XX
DR WPI; 1999-080907/07.
XX
DR P-PSDB; AAW94840.
XX
PT Recombinant hyperthermostable protease from Pyrococcus furiosus - and
XX gene encoding it, for large scale production of the protease for
XX industrial use.
XX
PS Disclosure; Page 53-54; 82pp; Japanese.
XX
CC The invention relates to a hyperthermostable protease derived from a
XX thermophilic bacterium (especially Pyrococcus furiosus). The protease has
XX working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
XX (optimum 6-8), and retains more than 90% of its activity after 8 hours at
XX 95 deg.C. The invention also provides gene sequences encoding a
XX polypeptide of formula SIG-Ala-Gly-Gly-Asn-PRO, where SIG is a signal
XX peptide from subtilisin, and PRO is the above protease. Host cells
XX (especially Bacillus strains) transformed with vectors comprising the
XX genes are used for the recombinant production of the protease. The
XX CC hyperthermostable protease which can be prepared in quantity suitable for
XX industrial use, can be used as an additive for drugs, washing agents and
XX foodstuffs and for chemical synthesis
XX
SQ Sequence 1977 BP; 453 A; 659 C; 554 G; 311 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,326-27 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.88 Conservative: 90
Best Local Similarity: 27.28 Mismatches: 203
Query Match: 14.64 Indels: 156
DB: 2 Gaps: 25
US-10-784-870-4 (1-639) x AAK05926 (1-1977)

```

QY	62	ValIGluse	rdIa	asnVal	lysLeu	eu---	LysGI	YLeu	eu	lysLeu	GI	uTh	Val	80											
													80												
Db	103	GTTCAG	CAAGAA	AACTCA	CGAC	CTCG	CAAC	CCCGG	CACT	GTTC	CAAGAA	AGTCC	AGAAGATG	162											
QY	81	ProAla	asnAsn	lysLeu	HisIle	ValGln	Phe	asnGI	ProIle	Leu	GI	uTh	lys	100											
													100												
Db	163	AACTGGA	ACCAGGA	AGATG	GAC	ACCCG	CTCAT	TAATG	TGGG	-----	-----	-----	201												
QY	101	Gln	lysLeu	GI	uTh	Thr	GI	Val	lysIle	Leu	asnP	TyrIle	Pro	asnP	Tyr	Val	120								
																120									
Db	202	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	207								
QY	121	Val	GI	uTh	Yr	GI	uTh	Val	asnP	Val	Gln	ser	lysVal	arg	ser	Ile	Gln	His	Val	140					
																			140						
Db	208	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	255					
QY	141	Glu	Pro	Tyr	Leu	Pro	lys	Tyr	lysIle	asn	Pro	Gln	Leu	Phe	Thr	-----	-----	-----	-----	158					
																			158						
Db	256	GTC	AAAGT	ATC-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	309					
QY	159	Ser	Thr	Leu	Val	lysVal	Leu	Val	Leu	Val	asn	P	Thr	lysGln	-----	-----	-----	-----	-----	176					
																			176						
Db	310	GAC	TTCT	CGCT	CGT	ATCG	CGG	GCAT	ATG	ATG	ACA	CGG	GT	ACT	TCG	GT	TA	CA	CA	AGG	CTCG	369			
QY	177	Gln	Leu	asn	arg	GI	Yl	Ile	GI	uTh	Ile	Val	asn	GI	uTh	Val	Ala	ser	asn	asn	Val	His	Tyr	196	
																						196			
Db	370	GGC	ATTA	ATA	GGT	TAT	CA	CA	GAG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	390					
QY	197	Thr	Ala	lys	Pro	GI	uTh	Yr	lysVal	Met	asn	asn	Val	Ala	arg	GI	Yl	-----	-----	-----	-----	-----	214		
																			214						
Db	391	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	438					
QY	215	Ala	asn	Val	Ala	Gln	ser	ser	Tyr	GI	Yr	GI	uTh	Val	asn	Val	Ala	234							
																	234								
Db	439	GCC	ATAC	CGCT	CGT	GA	AACT	CCCT	CGG	CTA	CA	GGA	AGG	GGT	GGT	GGT	GC	ATC	GT	C	498				
QY	225	Asp	Thr	GI	Yr	Leu	asn	P	Thr	GI	Yr	arg	asn	asn	ser	ser	Met	His	Gln	Ala	Phe	arg	GI	Yr	254
																							254		
Db	499	GAT	AC	G	G	G	T	A	T	A	G	A	C	G	G	-----	-----	-----	-----	-----	5				

Db	964	-----GTTGACAGCAACGACAAATCGCAGCTTCTCCAGCAGGGGACCGAC	1011
Oy	417	LySApGjYA:gIIeLYsProAspValMetAlaProGlyThrTyrlleLeuSerAlaArg	436
Db	1012	GCGACGGAAAGGCTCAAGCCCGGAAGTGTCTCCGCCCGCTTGACATCATAGCCCCGGC	1071
Oy	437	SerSerIeuValaProAspSerSerPheTrpAlaAsnHisAspSerLysTyrlaTyrlMet	456
Db	1072	GCCAGC-----GGAAACGAGCATGGGACCCCGCATTAACGACTACTACCAAGCC	1122
Oy	457	GlyGlyThrSerMetAlaThrProIleValaIaGly---AsnValaAlaGluIeuArgIu	475
Db	1123	TCTGGAAACAGCATGGCCACCCCGCAGCTTTCGGGCGTGGCGCGCTCATCTCCAGGCC	1182
Oy	476	HisPheValLysAsnArgGlyIleThrProLys-----ProSerIeuLysValaIa	493
Db	1183	CAC-----CCGAGCTGGACCCCGGACAGGTAAAGCCGC	1218
Oy	494	LeuIleAlaGlyIaIa-----AlaAspValaGlyLeuGlyTyrl	505
Db	1219	CTCATTCGAGACCGCGCAATAGTCGCCCCCAAGAGATAGAGGAGCATGCTTACGGTGC	1278
Oy	506	ProAsnGlyAsnGlnGlyTrpGlyArgValThrIeuAspLysSerIeu-----Asn	522
Db	1279	-----GGTAAGGTGAACCTCTACCAAGGCCATCAAGTACGACGAC	1317
Oy	523	ValAlaTyrlaAsnGluSerSerAlaIeuSerThrSerGlnLysAlaThrTyrlThrPhe	542
Db	1318	TACGCCAAGCTCACCTTTCACCGGCTCGTCCGCCCAAGGAGAGCGCACCCACCACTTC	1377
Oy	543	ThrAlaThrAlaGlyLysProLeuLysIleSerIeuValTrpSerAspAlaProAlaSer	562
Db	1378	GACGTCAGCGCGCACCTTGATGACGCGCACCTTACTACTGGAC-----	1422
Oy	563	ThrThrAlaSerValThrIeuValAsnAspIeuAspIeuValIleThrAlaProAsnGly	582
Db	1423	-----ACGGGCTCGAGCGCAATGCACCTTACTACCTACGACCCCAACGGG	1467
Oy	583	ThrArgTyrlValaGlyAsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArgAsn	602
Db	1468	AACGAG-----GTTACTACTCTTACACCGCCTTACTAC	1500
Oy	603	AsnValaGluAsnValPheIleAsnSerProGlnSerGlyThrTyrlIleGluValaGln	622
Db	1501	GGCTTCGAGGAAGTCGGCTACTACACCGCAGCGCGGAAGCTCGACGCTCAAGTGTCTC	1560
Oy	623	AlaTyrlAsnValProValGlyProGlnAsnPheSerIeuAlaIleValaAsn	639
Db	1561	AGCTACAAAG-----GGCGCGCGCAATACGAGTGCAGTCGTCTCAGC	1602
RESULT 13			
AAT85695			
ID	AAT85695 standard; DNA; 1962 BP.		
XX	AC	AAT85695;	
XX	XX		
DT	17-OCT-2003	(revised)	
DT	20-APR-1998	(first entry)	
XX	DE	Pyrococcus furiosus PFUS protease coding sequence.	
XX	KW	Protease; research reagent; thermal stability; pyrococcus furiosus; ss.	
XX	OS	Pyrococcus furiosus; DSM-3638.	
XX	PN	MO9721823-A1.	
XX	PD	19-JUN-1997.	
XX	PF	07-NOV-1996; 96MO-JP003253.	
XX	PR	12-DEC-1995; 95JP-00323285.	

QY 577 ILeThraLaProAnaGlyThraGlyValGlyAnaPheSerAlaProPheAspAn 596
DB 1450 CTCTACGATCCCAATGMAACAG-----GTTGACTAC 1482
QY 597 AenThraPheGlyAgaAenAnaValGluAnaValPheLeuAnaSerProGlnSerGlyThr 616
DB 1483 TCTTACACCGCGCTCATGATTCGAAAGAGTTGGTTATTACACCACGATGAGAAC 1542
QY 617 TyrThrIleGluValGlnAlaValThrAnaValProValGlyProGlnAnaPheSerLeuAla 636
DB 1543 TGGACATTAAGTTGTTAGCTACAGC-----GGAGTGCACCACTATCAAGTAGAT 1593
QY 637 ILeValAna 639
DB 1594 GTGGTAAGT 1602
RESULT 14
AAK05929
ID AAK05929 standard; DNA; 1962 BP.
AC AAK05929;
DT 06-MAY-1999 (first entry)
DE Hyperthermostable protease encoding DNA.
KM Hyperthermostable; protease; thermophilic; bacterium; subtilisin;
KW additive; drug; washing agent; foodstuff; chemical synthesis; ds.
XX Pyrococcus furiosus.
OS
PN MO9856926-A1.
PD 17-DEC-1998.
PF 04-JUN-1998; 98WO-JP002465.
PR 10-JUN-1997; 97JP-00151969.
PX 10-JUN-1997;
PA (TAKI) TAKARA SHUZO CO LTD.
XX Takakura H, Morishita M, Shimajo T, Asada K, Kato I;
XX WPI; 1999-080907/07.
DR P-PSDB; AAM94841.
XX
PT Recombinant hyperthermostable protease from Pyrococcus furiosus - and
PT gene encoding it, for large scale production of the protease for
PT industrial use.
PS Disclosure; Page 59-60; 82pp; Japanese.
XX
CC The invention relates to a hyperthermostable protease derived from a
CC thermophilic bacterium (especially Pyrococcus furiosus). The protease has
CC working temperature 40-110 deg.C (optimum 80-95 deg.C), working pH 5-10
CC (optimum 6-8), and retains more than 90% of its activity after 8 hours at
CC 95 deg.C. The invention also provides gene sequences encoding a
CC polypeptide of formula SIG-Ala-Gly-Asn-PRO, where SIG is a signal
CC peptide from subtilisin, and PRO is the above protease. Host cells
CC (especially Bacillus strains) transformed with vectors comprising the
CC genes are used for the recombinant production of the protease. The
CC hyperthermostable protease which can be prepared in quantity suitable for
CC industrial use, can be used as an additive for drugs, washing agents and
CC foodstuffs and for chemical synthesis
SQ
XQ Sequence 1962 BP; 602 A; 399 C; 471 G; 490 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,41e-24 Length: 1962
Score: 443.00 Matches: 163
Percent Similarity: 40.3% Conservative: 88
Best Local Similarity: 26.2% Mismatches: 204
Query Match: 13.6% Indels: 168

DB: 2 Gaps: 25
US-10-784-870-4 (1-639) x AAK05929 (1-1962)
QY 62 ValGluSerGluAnaValLysLeuLeuLysGlyLeuLeuLysLysLeuGluThrValPro 81
DB 103 GTTGAGAGAAACATATGCTGCTGCTACCGCAGACTGTTCAAGAAATATCAAAATTTGAAT 162
QY 82 AlaAnaAnaLysLeuHisIleValGlnPheAnaGlyProIleLeuGluGluThrLysGln 101
DB 163 CCTTACACAGAAATACAGACAGTA-----ATTGTAATTTGAAACCATAGCGAA 210
QY 102 Lys-----LeuGluThrThrGlyAlaValLysLeuLeuAspTyrIlePro 115
DB 211 AAAGAAATTCAGTACAGATTCGTTGAGTTATAGCGTCAAAAGTT----- 255
QY 116 AspTyrAlaTyrIleValGluTyrGluGlyAspValGlnSerLysValArgSerIleGlu 135
DB 256 -----AGTATATG----- 264
QY 136 HisValGluSerValGluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThr 155
DB 265 -----TACCATATTATACCCGCAATV----- 285
QY 156 LysGlyLaseThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAnaAnaLysGlu 175
DB 286 ---GCTGCCATCTTAAGTTAGAGACTTACTAGTCTCAGAGTTTACAGAGCGGTAA 342
QY 176 ValGlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaLaseAnaPheValHisTyr 195
DB 343 GCTAAGCTTCAGGTGTAGTTATTC-----CAGGAAAGCTACAA 384
QY 196 IleThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIleValLysAla 215
DB 385 GTTACAGTTTCAGCAGAAATTAGAGAGACTGAGTAGAGTCCACTCAAGTTATGGCAACT 444
QY 216 AspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAsp 235
DB 445 TACGTTTGAACCTTGGATATGAT-----GGTTTGGAAATCCAAATAGAAATATGAC 498
QY 236 ThrGlyLeuAspThrGlyArgAnaPheSerSerMetHisGluAlaPheArgGlyLysIle 255
DB 499 ACTGGAATTGAC-----GCTTCATCCAGATCCCAAGGAAAGTA 540
QY 256 ThrAlaLeuTyrAlaLeuGlyArgThrAnaAnaLys-----Asp 269
DB 541 -----ATTGGGTGGTAGATTGTCTCAATGTAGAGATTATCCATAGAT 585
QY 270 ThrAnaGlyHisGlyThrHisValAlaGlySerValLeuGlyAnaGlyAlaThrAsn--- 288
DB 586 GACCATGACATGGAATCTCATGAGCTTCAATAGCACAGCTGTACGAGCAAGCAATAT 645
QY 289 -----LysGlyMetAlaProGlnAlaAnaLeuValPheGlnSerIleMet----- 303
DB 646 GGCAGATGACAGGAAATGCTCCAGAGGCTTACCTGGCGGAAATTAAAGTTCTTAGGCGC 705
QY 304 AspSerSerGlyGlyLeuGlyLysLeuProSerAnaLeuGlnThrLeuPheSerGlnAla 323
DB 706 GATGGTTCTGGAAGCAATCTACTATATTAAAGGAGTTAGGCGCGGTGATTAACAA 765
QY 324 PheSerAlaGlyAlaArgIleHisThrAnaSerTyrGlyAla----- 337
DB 766 GATTAAGTACGAATTAAGATTAATCTTCTTGTGTTCAAGCCAGAGCTCATAGTGT 835
QY 338 -----AlaValAnaGlyAlaTyrThrThrAspSerArgAnaValAsp 351
DB 826 ACTGACGCTTAAGTACAGCTGTTAATGACGCGGATGCT----- 867
QY 352 AspTyrValArgLysAnaPheMetThrIleLeuPheAlaAlaGlyAnaGlyLysProAna 371
DB 868 -----GGATTAAGTTGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
QY 372 GlyGlyThrIleSerAlaProGlyThrAlaLysAnaAlaIleThrValGlyAlaThrGlu 391

```
Db 910 AAGTATACATCGGTTCTCCAGACGCTGCAAGCAAGTATTATACAGTTGAGCC----- 963
Qy 392 AATLeuabgProSerPhegLySerTyrAlaaphanlleaanhleValAlaaglPheSer 411
Db 964 -----GTTGACAGATGATGATGTTATTAACAGCTTCA 996
Qy 412 SerArgLyProThrLySerPglYArgIleValProaphValMetAlaProGlyThrTyr 431
Db 997 AGCAGAGGCCCACTGACAGCGGCGCTTAAGCTGAGGTGTGCTCCAGGAAACTGG 1056
Qy 432 lleuSerAlaargSerSerleuAlaProaphSerPheThrAlaanhleAaspSer 451
Db 1057 ATATATGCTGCGAGAGCAAGT-----GGAACCTAGCATGGGTCAACCAATTATATGAC 1107
Qy 452 lvsTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAlaaglYaaenValAla 471
Db 1108 TATTACACAGCAGCTCTCGGACATCATATGCAACTCTCAGCGTATGTCAGCC 1167
Qy 472 GluLeuabgIuhlsPheVallysaenArgLyIleThrProlys-----ProSerLeu 489
Db 1168 CTCTTCTCTCA-----GCAACCCGAGCTGAGCTCAGACAA 1206
Qy 490 LeuYaaAlaAlaLeuIleAlaGlyAla-----AlaAapVal 501
Db 1207 GTAATAACAGCCCTCATAGAACTGCTGATATGCTTAAGCCAGATGAAATAGCCGATATA 1266
Qy 502 GlyLeuGlyTyrProabngIyaenngIyTyrGlyArgValThrleuaplySerleu 521
Db 1267 GCCTACGGTCA-----GTAAGGTTATGATCATGACAGGCTATA 1305
Qy 522 AenValAlaTyrValaenGlyuserSerAlaLeuSerThrSerIlyVala----- 538
Db 1306 AAC-----TACATATACATAGTCAAGCTAGTTCATGATATGTTGCCAACAAAGGC 1359
Qy 539 -----ThrTyrThrPheThrAlaThrAlaGlyLysProleuYaaIleSerleuValTyr 556
Db 1360 AGCGAAACTCACAGATTCTGTTATTAAGCGAGCTCTGCTGTAACGCCACATTATATCTGG 1419
Qy 557 SerAapAlaProAlaSerThrThrAlaSerValThrleuValaAenAapleuAapleuVal 576
Db 1420 GACAAATGCCAAT-----AGCACCTTGATCTTTAC 1449
Qy 577 lleThrAlaProabngIyThargTyrValGlyAaenAapPheSerAlaProleuAapAa 596
Db 1450 CTCTACGATCCCAATGAAACAG-----GTTGACTAC 1482
Qy 597 AenTPAapGlyArgAaenAenValGluAenValPheIleAenSerProGlnSerGlyThr 616
Db 1483 TCTTACACCGGCTACATGATTCGAAAGGTTGGTTATTAACCAACCACTATATGAAACA 1542
Qy 617 TyrTyrIleGluValGlnAlaTyrAaenValProValGlyProGlnAapPheSerleuAla 636
Db 1543 TCGACAAATTAAGGTGTATAGCTACAGC-----GGAAGTGCANAACTATCAAGTAAAT 1593
Qy 637 lleValaen 639
Db 1594 GTGTAAAT 1602
RESULT 15
ACT64496/c
ID ACT64496 standard, DNA; 9125 BP.
XX ACT64496;
XX 02-JUN-2005 (first entry)
XX M. xanthus DNA fragment, seq id 959.
XX Transgenic plant; DNA replication; gene regulation; gene expression; de.
XX Myxococcus xanthus.
XX
```

```
PN US6833447-B1.
XX
XX 21-DEC-2004.
XX
XX 10-JUL-2001; 2001US-00902540.
XX
XX 10-JUL-2000; 2000US-0217863P.
XX
XX (MONS ) MONSANTO TECHNOLOGY LLC.
XX
XX Goldman BS, Hinkle CJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX
XX
XX New substantially purified Myxococcus xanthus nucleic acid molecule
XX encoding a nitrite reductase, useful for determining gene expression,
XX identifying mutations in a gene of interest, and for constructing
XX mutations in a gene of interest.
XX
XX Example 1; SEQ ID NO 959; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
XX encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX recombinant DNA construct for expression of a nitrite reductase gene in a
XX plant cell, and a plant cell comprising the recombinant DNA construct.
XX The nucleic acid is useful for determining gene expression, identifying
XX mutations in a gene of interest, and for constructing mutations in a gene
XX of interest. Sequences given in records for SEQ IDs 1-1849 represent a
XX set of 1849 contig and singleton sequences comprising coding sequences,
XX CC DNA replication elements, promoters and other regulatory elements from
XX the genome of the bacterium Myxococcus xanthus. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO
XX
SQ Sequence 9125 BP; 1331 A; 2947 C; 3329 G; 1518 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 9.73e-24 Length: 9125
Score: 443.00 Matches: 172
Percent Similarity: 42.94 Conservative: 108
Best Local Similarity: 26.34 Mismatches: 231
Query Match: 13.64 Indels: 142
DB: 14 Gaps: 29
US-10-784-870-4 (1-639) X ACT64496 (1-9125)
Qy 60 PheLeuValGluSerGluAenValLyLeu-----LeuYaaGly-----leuLeu 74
Db 8329 TACTGTGTCGGGCTGAGACAGCGGCTTGCACCGGCTTAAGGGGCGACAGCCCTCGTGC 8270
Qy 75 lvsYaaLeuGluThrValProAlaAaenAenYaaIleValGlnPheAenGlyPro 94
Db 8269 GCGGGGAGTCTACACTCGGAGCGGACCAACCTGCACTGTGATTCACGCGCAG 8210
Qy 95 lleLeuGluGluThrLyseGlnLyseGluThrThrGlyAlaLyIleleuAapTyrIle 114
Db 8209 CCCCTGCGGAGTACCGGACCGCATCAAGCGCGGGGCGAAGTGTGCGCTTCTTC 8150
Qy 115 ProAapTyrAlaTyrIleValGluTyrGluGlyAapValGlnSerLyValAargSerIle 134
Db 8149 ACGACACACACTTCTGCTGAGATGATGCCAGACGACGACAGCAGACGCGCTCGGCTG 8090
Qy 135 GluIleValGluSerValGluProTyrleuProTyr----- 147
Db 8089 CCTATATGTCGCTGAGTGGGCGGTACATCCGAGTACCGGCTGGAAGGCGTCTGCC 8030
Qy 148 -----lysIleAapProGlnleuPheThr----- 155
Db 8029 GAGTCTTCAGAGGCGCTCGCTGAAAGCTGAGGCGGACGCGCTACTCATCATGTGGGCG 7970
Qy 156 ---lysGlyAla-----SerThrleuVallyAlaAlaLeuAap 168
Db 7969 GAGCGGCGCGGACAGACAGAGGCGAGGTGCGCTGCTGTCGACCGCTGGGCGGACG 7910
```

QY 169 Thrlysglnaenvalygluvalgluleuarglyle-----181
 Db 7909 GTGGAGCTC-----ATCGAAGCGGGCGGCTCGCGGTGAGGCCACCTCCACC 7862
 QY 182 ---gluileuileaglityvalalaserbnaerValhlytyllethr-----Ala 198
 Db 7861 CAGTCCCAAGTGGAGCGGTGATTCGCTCCAAAGCGCGTGCAGTACGACCGCTGGGGC 7802
 QY 199 lueploglytylvalvalmetabnaerValalaglyllevallyalabryValala 218
 Db 7801 GGGCGGGGT---GAGCTCGCAACAACAAGCTCGCGATGTGGGGCGGCGCACTTACCTC 7745
 QY 219 GlnserseTyrglyleuTyrglynglynglyleValalavalalaserbnaerVal 238
 Db 7744 GAAGGCGTACAGAGCTTCAGCGGACAGGGCGTCCCGCGGAATCTTCGACACGGAAGTGG 7685
 QY 239 Asprthrglyargbnaer-----SerSerMetHla 249
 Db 7684 CGCACCGACGACAGAGTGGCGGACACCTCCCTCATCCACAGACGTCACACGAGGAC 7625
 QY 250 AlaerhnaerglylyllethrAlaleuTyrglyalaleuTyrglyThrAsbnaerVal 269
 Db 7624 ACCTTCACAGGT---ACCGCTGCTACAGCATC-----7595
 QY 270 ThranglyhlygllyThrhlvalalaglyserValleuglyValaglyalathrbnaer 289
 Db 7594 ---AACTTCGCGCGGGCGGTGACGGGAACCG---CGC 7562
 QY 290 GlyMetAlaProglhAlabnaerValrheglnserlleMetAspSerseTyrglyleu 309
 Db 7561 GGCCTGGTCCCTCCGACAAAGCATCTTACCGGTACACAGAGTCTCC---CAGTTC 7505
 QY 310 GlyglyleuProserbnaerleu-----GlnThrleuPheSerGlnAlaPheSerAla 326
 Db 7504 GGGGGGCGCAAGTCTCGTACACATCAACCGGAGCTACGACCGCGCGCGGTAC 7445
 QY 327 GlyAlaargylehlythrAsbnaerTyrglyalalalavalaglyalatyThrThrAsp 346
 Db 7444 CGGCGGTGTTCAGACGTCAGCGGTGGGACGCGCGAGTGCAGCTACACACCATC 7385
 QY 347 SerArgbnaerValAspAsprTyrglyValarglybnaerMetThrleuPheAlaagly 366
 Db 7384 TCGCGAGAGTGTATGATCTGTTCCAGTACCC---ATCCTCAGCACGCGAG 7334
 QY 367 AsnglyuargProbnaerglylyThrIleSerAlaPro---GlyThrAlaTybnaerAla 385
 Db 7333 TCACAGACGACACCGGCAACGAACTCCGCGCACAGCGGTGGGGAAGAACATCGTG 7274
 QY 386 ThrValglyAlaThrGluabnaerArgProSerPheglySerTy---AlaAspAla 404
 Db 7273 TCCGTGGGT-----GGCATGTACCACTACGACACACCC 7241
 QY 405 AsnhleValAlaGln-----PheSerSerArglyProThrlybnaerGlyArg 420
 Db 7240 AACCGCGGAGTACCGGTGAAATGGCGGCGCGACGCTGGGTCCGGCGGCGAGCGCGG 7181
 QY 421 IlelyProbnaerValMetAlaProglyThrTyrlleuSerAlaargSerleuAla 440
 Db 7180 CTCAGCGCGGACGCTGTACTACGACCTCATCGCTCGCGGTGAAACACCC-----7127
 QY 441 ProAspSerPheThrAlaAsnhleAspSerlyTyrglyAlatyMetglylyThrSer 460
 Db 7126 ---AGCAACACGTCCTACACCAACTTCGGCGGACCGAC 7091
 QY 461 MetAlaThrProleValAlaglybnaerValalaglyleuarglyhlyspheVallybna 480
 Db 7090 GCGGCGACGCGCGACAGCGGGGCACTTGGGCTGTCTTCAGAGTGG-----CAC 7037
 QY 481 Argglylle-----ThrProlybPro 487
 Db 7036 GAGGCGCTGTGGCGGCTTCGGTGTGGCGGAGTCTTCGACAGCGCTCGAAGATG 6977

QY 488 SerleuLeuValAlaAlaLeuIle-----AlaGlyAla 498
 Db 6976 GCCACGGCGAAGCGCGGTGATGATCAATGAGCCCATCGTCAACATCGCTCGCGGCGCGC 6917
 QY 499 AlaAspVallyleuglyTyrglyProbnaerGlybnaerGlyTyrglyArgValThrleu 517
 Db 6916 TCCAGCGGTACCTG-----ACGCGCGCGCGGACAGGCGTGGGTACCGCGGACGTGAAG 6863
 QY 518 ---AspySerleuAsnValAlatyValabnglyserSerAlaLeuSerThr 534
 Db 6862 CGCGTGTATACCGCGCGCGCGGTGACAGCATCATCAACAGACGCGACCTGCTCCG 6803
 QY 535 SerGlnlyAlaThrTyrrPheThrAlaThrAlaglylys---ProleuTyrlleSer 553
 Db 6802 CTCGCGCAAGATACGTCAACGTCAACGCGGTCTCCGCGGAGACGAGCTCAACGTACC 6743
 QY 554 leuValTrpSerAspAlaProAlaSerThrThrAlaSerValThrleuValaAspLeu 573
 Db 6742 ATGATCTACACGACCGCGCGGACCGGTGGCGCGGCGGACCGCGCATCAACGACCTG 6683
 QY 574 AspleuValilleThrAlaProbnaerGlyThrArgTyrglyValbnaer---PheSerAla 592
 Db 6682 TCGCTGCGCGGTACGTCCCGCACCGCGGTGCTTACTGGGGCAACAACGCGCTGACGGCG 6623
 QY 593 ProPheAspAsnAlaTrpAspGlyArgAspAsn-----ValGluAsnValPhe 608
 Db 6622 AGCAACGTCCTCACGGCGGCGCGGTGTGAAACAAGGTGATACGCTGAGAACGTCTTC 6563
 QY 609 IleAsnSerProGlnserGlyThrTyrlleGlyVal 621
 Db 6562 CTCGCAACCGCGCGCGGCGGACGTGACCGTGAAGGTG 6524

Search completed: April 8, 2006, 03:44:56
 Job time : 1707 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 03:09:21 ; Search time 6632 Seconds
(without alignments)
4507.983 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267

Sequence: 1 MRRKKVPLSLVSAALSLTV.....EVOANVPVGPONSLATVN 639

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.model -DEV=xlh
-O=/abs/ABSSMB.spool/US10784870/runat_07042006.143417_27495/app.query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEADS=500 -MINLEN=0 -MAXLEN=200000000 -HOST=adbb08
-USER=US10784870 @CGN 1 1 5315 @runat_07042006.143417_27495 -NCPU=6 -ICPU=3
-NO NMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_gest1.*
10: gb_gest2.*
11: gb_gest3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	282.5	8.6	689	8	CV960494 PKRXC 89
2	258	7.9	1605	5	BQ622771 CC Contig
3	246.5	7.5	627	10	CV960494 PKRXC 89
4	245.5	7.5	4398	4	CR861261 Pongo pyg
5	244	7.5	4198	4	AK029048 Mus muscu
6	243	7.4	3159	11	DD050980 Homo sapi
7	243	7.4	3159	11	DD050981 Pan trogl

8	242.5	7.4	679	10	CV960494	689 bp	mRNA	linear	EST 25-JAN-2005
9	231	7.1	540	3	BQ622771				
10	220.5	6.7	594	3	BQ622771				
11	219.5	6.7	1595	4	CV960494				
12	219	6.7	962	8	DR651450				
13	217.5	6.7	532	11	TA319610P				
14	217.5	6.7	874	7	CV960494				
15	214	6.6	555	10	CV960494				
16	214	6.6	729	5	BM449844				
17	213.5	6.5	839	7	CV960494				
18	213.5	6.5	863	7	CV960494				
19	213.5	6.5	2337	10	CV960494				
20	212	6.5	632	10	CV960494				
21	212	6.5	2619	10	CV960494				
22	209	6.4	841	7	CV960494				
23	208.5	6.4	2141	3	BQ622771				
24	208	6.4	2208	5	CV960494				
25	207	6.3	720	5	BM449844				
26	205.5	6.3	826	7	CV960494				
27	202.5	6.2	2202	10	CV960494				
28	202	6.2	939	8	DR602738				
29	201.5	6.2	1771	4	CV960494				
30	201	6.2	508	9	CV960494				
31	200	6.1	1550	8	DR010221				
32	199.5	6.1	601	3	BQ622771				
33	199.5	6.1	744	8	DR628193				
34	198.5	6.1	707	8	DR628018				
35	198.5	6.1	885	8	DR624615				
36	197	6.1	797	8	DR657363				
37	197	6.0	822	8	DR657362				
38	196	6.0	633	3	BQ622771				
39	196	6.0	690	8	DR624989				
40	195	6.0	771	6	CA320325				
41	195	6.0	1002	11	CV960494				
42	194.5	6.0	2274	10	CV960494				
43	193.5	5.9	824	7	CV960494				
44	193	5.9	2472	4	CV960494				
45	192.5	5.9	659	5	BM448603				

ALIGNMENTS

RESULT 1
CV960494 689 bp mRNA linear EST 25-JAN-2005
PKRXC 8939 mycelium, starved in water Phytophthora infestans cDNA,
DEFINITION
CV960494
ACCESSION
CV960494.1 GI:58150285
VERSION
CV960494.1
KEYWORDS
SOURCE
ORGANISM
Phytophthora infestans (potato late blight agent)
Phytophthora infestans
Bukaryota; stramenopiles; Oomycetes; Pyliales; Pylaceae;
Phytophthora.

REFERENCE
1 (bases 1 to 689)
Randall,T., Dwyer,R.A., Huitema,E., Beyer,K., Cvitanich,C.,
Kelkar,H., Fong,A.M., Gates,K., Roberts,S., Yaczkan,E., Gaffney,T.,
Law,M., Testa,A., Torto-Alalibo,A., Zhang,M., Zheng,L., Mueller,E.,
Windass,J., Binder,A., Birch,P.R.J., Gisel,U., Govers,P., Gow,N.A.,
Mauch,P., van West,P., Maugh,M.E., Yu,J., Bolter,T., Kamoun,S.,
Lam,S.T., and Udelsdon, H.S.
Large-scale gene discovery in the oomycete Phytophthora infestans
reveals likely components of phytopathogenicity shared with true
fungi

JOURNAL
PLUMED
COMMENT
MOL. Plant-Microbe Interact. 18 (3), 229-243 (2005)
15782637
Contact: Judeleison HS
Department of Plant Pathology
University of California
Weber Hall, Riverside, CA 92521, USA
Tel: 909 787 4199
Fax: 909 787 4294
Email: howard.judeleison@ucr.edu.

FEATURES
source
Location/Qualifiers
1..689
/organism="Phytophthora infestans"
/mol_type="mRNA"
/strain="88069"
/db_xref="taxon:4787"
/sex="A1"
/clone_lib="mycelium, starved in water"
/note="Vector: pSPORT1"

ORIGIN

Alignment Scores:
Pred. No.: 2,19e-20 Length: 689
Score: 282.50 Matches: 84
Percent Similarity: 51.3% Conservative: 34
Best Local Similarity: 36.5% Mismatches: 79
Query Match: 8.6% Indels: 33
Gaps: 10

US-10-784-870-4 (1-639) x CV960494 (1-689)

QY 231 VALAVALAALAPThrGlyLeuAPThrGly-ArgAAsp-SerSerMetHis---- 248
DB 3 AATGGGAAAGCTGTACGCTGACGGTACGGTCCGGATTCGCGGTCGACCGCGCT 62
QY 249 -----GluAlaPheArgGlyLeuIleThrAlaLeuTyAlaLeuGlyArgThraAana 267
DB 63 CCGCGGACCGGTGGCG 119
QY 267 laAAsp-----ThraAnglyAlaGlyThraIleValAlaGlySerValLeuG 283
DB 120 CGGATGAAGCCGAAACTATCACTTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
QY 283 lyaAnglyAlaIleThrLeuIle-----GlyMetAlaProGlnAlaIleLeuValP 299
DB 180 GAGGCTCGTTAGCGGTAAACATGCCAATCTTGCGCATTCGCGCAAGCGACGATGAT 239
QY 289 heGlnSerIleMetAAspSerGlyGlyLeuGlyGly----- 311
DB 240 TTATGACATCGGAGCCGACGCGGAAACCTGCACAGGTACGTCGATGTCGCGTCCAC 299
QY 312 -----LeuProSerAAspLeuGlnIleThrLeuPheSerGlnAlaIleSerAlaGlyAlaArgI 330
DB 300 TAGCCACACCTGCACG 359
QY 330 leHISThrAAspSerITrGlyAlaAlaValaAnglyAlaIleTyThrIleAAspSerAlaGlyAAsp 350
DB 360 TTTTCTCATTCCTCGTGGGCGACACCT--GGGTGGATTCAGCTCACAGCCAGAGATT 416
QY 360 aAlaAspTyAlaIleGlyAAsp---AAspMetThrIleLeuPheAlaIleGlyAAsp--- 367
DB 417 TGGAGCGCTTTATTTACGAAACCCAGACGTCCTGCTGCTGGCGCGCGCGCGCGCG 476
QY 368 --GluArgProAAspGlyGly-----ThrIleSerAlaProGlyThraIleValaAnglyAlaI 385
DB 477 GCGAGCTTCGACG 536
QY 385 leThrValGlyAlaIleThrGluAAspLeuAAspProSerPheGlySerTy-----A 401
DB 537 TTTTGGTGGCGCATCGCTGAACTGCACTCTTTTCAAGCATTTGGGTCGCGCGCG 596
QY 401 laAAspAAspIleAAspIleValaIleGlnPheSerSerArgGlyProThrIleAAspGlyAlaArgI 421
DB 597 TATCCAAAGAGCGACAGTGTGATGCTTTCTCTGACGGCGCCCAACGACGCGAGAT 656
QY 421 leIleProAAspValMetAlaProGly 429
DB 657 TAAAGCCGATGTAGTGGCGCGGT 682

RESULT 2
B0622771 1605 bp mRNA linear EST 01-JUL-2002
LOCUS B0622771
DEFINITION CC_Config7 Condiolobolus cornatus ARSEF 512 Condiolobolus cornatus

ACCESSION
B0622771
VERSION
B0622771.1 GI:21649940
KEYWORDS
EST
SOURCE
Condiolobolus coronatus (Delacroixia coronata)
ORGANISM
Condiolobolus coronatus
Eukaryota; Fungi; Zygomycota; Zygomycetes; Entomophthorales;
Ancylistaceae; Condiolobus.
REFERENCE
1 (bases 1 to 1605)
Freimoser, F.M., Screen, S., Hu, G. and St. Leger, R.J.
EST analysis of genes expressed by the zygomycete pathogen
Condiolobolus coronatus during optimized secretion of proteins
Unpublished (2002)
JOURNAL
Contact: Freimoser F. M.
Department of Entomology
University of Maryland
412 Plant Sciences Building, College Park, MD 20742, USA
Tel: 301 405 16 13
Fax: 301 314 92 90
Email: ff34@umail.umd.edu.

FEATURES
source
Location/Qualifiers
1..1605
/organism="Condiolobolus coronatus"
/mol_type="mRNA"
/strain="ARSEF 512"
/db_xref="taxon:34488"
/clone_lib="Condiolobolus cornatus ARSEF 512"
/note="Vector: Unizap; Condiolobolus coronatus was grown in minimal medium supplemented with Manduca sexta cuticle and perone for 18 hours. A cDNA library was constructed in the unidirectional lambda vector Unizap."

ORIGIN

Alignment Scores:
Pred. No.: 4.67e-17 Length: 1605
Score: 258.00 Matches: 108
Percent Similarity: 43.1% Conservative: 58
Best Local Similarity: 28.1% Mismatches: 111
Query Match: 7.9% Indels: 108
Gaps: 22

US-10-784-870-4 (1-639) x B0622771 (1-1605)

QY 119 TyrIleValGluTyArgGlyAAspValGlnSerTyValaArgSerIleGlnIleValGlu 138
DB 514 TATATCTTTTTCGATCAAAATGCTGTCAGCTATGATGCT-----CAAGTGA 567
QY 139 SerValGluProTyLeuProTyLeuTyIleAAspProGlnLeuPheThr-----Lys 156
DB 568 TCTGTT-----ACCTCTCTTCACGACCGGAT 597
QY 157 GlyAlaSerThrLeuValIleValaLeuAlaLeuAAspThrTyGlnAAspValGlyVal 176
DB 598 GGTTCACACGCTCTTCAAC-----AAATTCATCCGCTTCTACGCTGTCGCTGT 651
QY 177 GlnLeuArgGlyIleGlnGlnIleAlaGlnTyValaIleSerAAspValIleTyIle 196
DB 652 AAGCTCAAT-----GATGAACCTTGAAAGTCAAGCT-----CTTTCACGTC 699
QY 197 ThrAlaAspProGluTyTyValaIleAAspValAlaArgGlyIleValaIleAAsp 216
DB 700 ACC-----AAGTTGAACAAGAT-----GGTGTGCTGAAGCTATT 735
QY 217 ValaIleGlnSerSer-----TyGlyLeu----- 224
DB 736 GCCGTCCAGGCAATGCTCTTGGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
QY 225 -----TyGlyGlnGlyGlnIleValaIleValaIle 234
DB 796 TCTGCTCTTACTTTTCAACCATGCTGATGCTGATCAAGGTGCTGCTGCTGCTGCTT 855
QY 235 AAspThrGlyLeuAAspThrGlyArgAAspSerSerMetHisGlnAlaPheArgGlyLys 254

Db 856 GATACCTGCTGTCATATGTCAGCCACAAATGAC-----TTGGTGGTGGT 897
Qy 255 TTTTThAlaLeuTyAlaLeuGlyArgTrpAsnAla-----AsnAspTrp 270
Db 898 GCCACT-----TGGGGTACTACACAGCTGGTGTAGCAACACTGATGCT 942
Qy 271 AaAGLYHISGLYThrHisValAlaAGLYSerValLeuGlyAaAGLYAlaThrAsnLYSGLY 290
Db 943 CACGGTCACGGTACTAGCTAGCTGGTGGTACTATTGCTGGT-----ACCACCTAATGGT 993
Qy 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
Db 994 GTTGGCAAGAGAGCTAACTATTGTTGGCGTTAAGCTGTATGATATGGCTCCGATCA 1053
Qy 309 LeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAla 328
Db 1054 TACTCTGGAAATTAATCTCCCGTATTTGACTGGGTTGTAAACAC-----TCTGCTGCAAG 1107
Qy 329 ArgIleHisThrAsnSerTrpGlyAlaAlaValAaAGLYAlaTyThrTrpAspSerArg 348
Db 1108 AAAGTATCTATATGAGTTTATGAGGTGTAGAGGTGTAAAGACGATGCTTTAACT----- 1158
Qy 349 AsnValAspAspTrpValArgLYAsnAspMetThrIleLeuPheAlaAlaAGLYAaAGLY 368
Db 1159 GCTGTTAAACAGCTCTCAGCAAGGAGTGTACACTGTCTC---CTGCTGTAAACGAT 1215
Qy 369 ArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLYAsnAlaIleThrValGly 388
Db 1216 AACAGAGATGCTTGGATATCTCT---CCGCTTGTGCTCCCTCCGCAATTAACGTTGGAT 1272
Qy 389 AlaThrGlu-----AsnLeuArgProSerPheGlySerTrpAlaAspAsnIleAsnHis 406
Db 1273 GCCACTGAAGTCAATATAAAAGGCTCAATCTCACTCGTAGCTGTGCT----- 1326
Qy 407 ValAlaGlnPheSerSerArgGlyProThrLYAspGlyArgIleLYsProAspValMet 426
Db 1327 -----GATTACTTA 1335
Qy 427 AlaProGlyThrTrpIleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheTrp 446
Db 1336 GCTCTGGTGTCAACATTCCTCTCCACC-----TGG 1365
Qy 447 AlaAsnHisAspSerSerTrpAlaTyTrpMetGlyGlyThrSerMetAlaThrProIleVal 466
Db 1366 AAGGATCTTAACACTGCCAACACCAACCACTCTGTGTAACCTTAATGGCTTGCCTCAATT 1425
Qy 467 AlaGlyAsnValAla 471
Db 1426 GCTGGTTAGCTGCT 1440

RESULT 3
LOCUS CW807544 627 bp DNA linear GSS 20-JUN-2005
DEFINITION pine 1964 Phytophthora infestans T30-4 phytophthora infestans
genomic, genomic survey sequence.
ACCESSION CW807544
VERSION CW807544.1 GI:57981961
KEYWORDS GSS.
SOURCE Phytophthora infestans (potato late blight agent)
ORGANISM Phytophthora infestans
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.
1 (bases 1 to 627)
Randal, T.A., Dwyer, R.A., Huitema, B., Beyer, K., Cvitanich, C.,
Kelkar, H., Fong, A.M., Gates, K., Roberts, S., Yatskan, B., Gaffney, T.,
Lew, M., Testa, A., Torto-Alalibo, T., Zhang, M., Zheng, L., Whelley, B.,
Windase, J., Binder, A., Birch, P., Gisi, U., Goyers, F., Gow, N.A.,
Mauch, F., van West, P., Waugh, M.B., Yu, J., Bolter, T., Kamoun, S.,
Lam, S.T. and Judelson, H.S.
Large-scale gene discovery in the oomycete Phytophthora infestans
reveals likely components of phytopathogenicity shared with true
fungi
Mol. Plant-Microbe Interact. 18 (3), 229-243 (2005)

PUMED 15782637
COMMENT Contact: Judelson HS
Department of Plant Pathology
University of California
PLPA, Riverside, CA 92521, USA
Tel: 9518274199
Fax: 9518274294
Email: howard.judelson@ucr.edu
Class: Shocgun.
FEATURES
source Location/Qualifiers
1..627
/organism="Phytophthora infestans"
/mol_type="genomic DNA"
/isolate="T30-4"
/db_xref="taxon:4787"
/sex="A1"
/clone_lib="Phytophthora infestans T30-4"
/note="Vector: pUC18"

ORIGIN
Alignment Scores:
Pred. No.: 1,94e-16 Length: 627
Score: 246.50 Matches: 77
Percent Similarity: 46.8% Conservative: 24
Best Local Similarity: 35.6% Mismatches: 72
Query Match: 7.5% Indels: 43
DB: 10 Gaps: 8
US-10-784-870-4 (1-639) x CW807544 (1-627)

Qy 269 AspTrpAsnGlyHisGlyThrHisValAlaAGLYSerValLeuGlyAaAGLYAlaThrAsn 288
Db 12 GATTCACACGTCACAGTACATTCGTTACGGGAATTATGAGGCGCA-----TTGGGC 65
Qy 289 LYSGLYMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyGly 308
Db 66 TTGGAGGTGGCACCGGAGACTCTGATGGCATGTAAAGATGTCTGCTCAAGGTTC 125
Qy 309 LeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe----- 320
Db 126 TATGGG-----TGGATTTGGTACAGATTTTCCAGTCAATGTGTGCGCTACAGCA 176
Qy 321 -----SerGlnAlaPheSerAlaGlyAlaGlyIleHisThrAsnSerTrpGly 336
Db 177 CCAAGTGGACATCAAAAGACTGCTCCAAAGCAGCAAGAAATGTAAACAGTTGGCGGT 236
Qy 337 AlaAlaValaAsnGlyAlaTyThrTrpAspSerArgAsnValAspAspTrpValArgLYs 356
Db 237 -----GGCGGTCAAGCTCTCAAGATGTTGATGGGGTTATCAACGCTGGCGAGCT 287
Qy 357 AsnAspMetThrIleLeuPheAlaAlaAGLYAaAGLYArgProAsnGlyGlyThrIleSer 376
Db 288 GCAAGGATCAATCCCTGCTAGCTCAGGAAATACGTGCTTAATCGCGTACTATTGCA 347
Qy 377 AlaProGlyThrAlaLYAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSer 396
Db 348 TCGCTGGAGATTCGGCTAGTGTGATCAACGTAAGTGGCAACCGAT----- 392
Qy 397 PheGlySerTrpAlaAspAsnIleAsnHis-----ValAlaGlnPheSerSerArgGlyPro 415
Db 393 -----ATCAACGACGAGCTGCTCTTTACAGCAAGACCGCCG 431
Qy 416 ThrLYAspGlyArgIleLYsProAspValMetAlaProGlyThrTrpIleLeuSerAla 435
Db 432 AAGAGTCAAGCTCTCCGAAACCGATGTGGACGCCACAGGCTCTGTTATG----- 485
Qy 436 ArgSerSerLeuAlaProAspSerSerPheTrpAlaAsnHisAspSerTrpAlaTyArg 455
Db 486 -----TCATCTGCTGAGAGGAC-----GAGCGCTTAACTGCTTC 521
Qy 456 MetGlyGlyThrSerMetAlaThrProIleValAlaAGLYAsnValAla 471

Oy		356	LysAsnAspMetThrIleLeuPheAlaIleValGlnGluArgProAsnGlyIleThyrIle	375
Db		1511	GCTACAAATGTAATCAAGGTTCCTGCTATTGGCAATGACGACCTCTTAAAGCACCTCG	1570
Oy		376	SerAlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgPro	395
Db		1571	AATAACCCGCATGAATAATGATGAATTTGGAGTAGTGCGCATTTGAC-----	1618
Oy		396	SerPheGlySerTyrlaPheAsnIleAsnHisValAlaGlnPheSerSerArgGly---	414
Db		1619	-----TTTGAAAGATTAACATC-----GCCCGCTTTCTTCAAGGGGAATG	1655
Oy		415	-----ProThrLysAsnProIleValGlyLeuProAsnValMetAlaProGly	429
Db		1658	ACTAAGCTGAGCACTAACAGAGCACTACGGTGGCAAGAAGCCCTGAATATGTC-----	1708
Oy		430	ThrTyrlleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheThrAlaAsnHis	449
Db		1709	ACCTATGGTGCCTGGCGTGGGGGTTCTGTGTGGAAAGGGGGGTGC-----	1755
Oy		450	AspSerLysTyrlaTyrlaTyrlaGlyIleThrSerMetAlaIleThrProIleValAlaGlyAsn	469
Db		1754	-----CGGGCCCTCTCAAGGACCAAGTGTGCTTCCAGTGGTTGCAAGTCT	1801
Oy		470	ValAlaGlnLeuArgGlnHisPheValLysAsnArgGlyIleThrProLysProSerLeu	489
Db		1802	GTACCTTATTAATGAGACACAGTCCAGAAAGCCGTGACGTGTG-----AATCTGGCCAGT	1855
Oy		490	LeuLysAlaAlaLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTyrlaProAsnGlyAsn	509
Db		1856	ATGAAGCAGGCCCCCTGATCCCGTCAAGCCCGAGAGCTC-----CCCgggGTCAAC	1903
Oy		510	-----GInglyTrpGlyArgValThrLeu-----	517
Db		1904	ATGTTTGAGCAAGGCCACAGGCAAGCTCATCTGCTCAAGAGCTTACAGTCTCAACAGC	1963
Oy		518	-----AspLysSerLeuAsnValAlaTyrlaValaGlnGluSerSerAla-----	531
Db		1964	TACAAGCCAAGGCAAGTTTGAAGCCCAAGCTACATGATGTGACAGAGTGTCCCTCAATG	2023
Oy		532	---LeuSerThrSerGlnLysAlaIleThyrTyrlaPheThrAlaThrAlaGlyLysProLeu	550
Db		2024	TGGCCCTACTGCTCCACGCCATCTACTAT-----GGGGGAATGCC---	2055
Oy		551	LysIleSerLeuValTrpSerAspAlaProAlaSerThrThrAlaSerValThrLeuVal	570
Db		2066	-----ACAGTTGTTAATGCACCATCTCTC	2089
Oy		571	AsnAspLeuAspLeu-----ValIleThrAlaProAsnGlyThyrArgTyrlaGly	587
Db		2090	AACGGCAATGGAGGTGCACAGAGAATTTGTATGATTAAGCTGACCTGACAGCCCTATATG--	2146
Oy		588	AsnAspPheSerAlaProPheAspAsnAsnTrpAspGlyArgAsnAsnValGluAsnVal	607
Db		2147	-----CCACAGAACGGAGAC-----AAcATTGAAGTCCGC	2176
Oy		608	PheIleAsnSer-----ProGlnSerGlyThyrTyrlaIleGluVal	621
Db		2177	TTCTCTTAATCTCTGTGCTTAAATGGCTTGTGTCAGGCTACCTGGCATCTCCATT	2230
RESULT 5				
AK029048				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
REFERENCE
AUTHORS
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Komoto, H., Aklyama, Y., Nishikawa, K., Kitamura, T., Tashiro, H., Itoh, M.,
Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaichiwa, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11076861
REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
PUBMED
11076861
REFERENCE
AUTHORS
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
PUBMED
12041101
REFERENCE
AUTHORS
6 (bases 1 to 4198)
Adechi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagata, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, F., Imokani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M.,
Koike, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N.,
Okazaki, Y., Satoh, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE
Direct Submission
JOURNAL
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp).
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .4198
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:4732484M11"
/db_xref="taxon:10090"
/clone="4732484M11"

```
/issue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
417..3575
/feature="unnamed protein product; membrane-bound
transcription factor protease, site 1 (MGD) [MG:1927235,
GRLNM_019709, evidence: BLASTN, 99%, match=3782]
putative"
/codon_start=1
/protein_id="BAC26263.1"
/db_xref="GI:26325018"
/translation="MKLVSTWLLVLLVLLCGKHLGDRLEKTRALAKAPCSHLTK
VEESTVVEVEYVAENVGYFTAKANSPISALKSEVENMRLIPNNNSQVPSDFE
VIOIKKOKAGLLTLEDHPNIRKTPORVPSIKAEENPIVPCNETWSQKQSSR
PIKASLSISGSGFWHATGHSRRLRLAPROYAQTLDVLMOKETANRVAAPD
TGLSEKHPKVKERTNMTNERTLDDGLGHGTPVAGVLAIVMCEGFPADLELIFR
VFTNNQVSYTSWFLDAFNALIKMDVLMISIGDPFMDHPMDKVELTANNVIMVS
ALGNDPLGYTLNPPADQMDVIGGIDFEDNIAKRSRGMNTWELPGGAYKVDIV
TYGAGVRSKGKGCALSGTSVSPVAGVATLVAIVTQKELVPAVKALIASA
RLPGVNMPEOGHGKLDLRAVOILSSYTPDASISPSYDILTECPMPMYSCQPIYTG
CMPTIVNVTILNMGYTKRIIVDKPERPILPONGDIEVAFTSSVLPMSGLIISI
SVTKAASWEGIAQGHIMTVASPAETELHSGAENSTVAKLIPKXILIPPRSRVL
WQYHNLRAPGVFPEDNLRKNDPLDNGDVHNTFRMDYHLSMGYFVLAAPF
TCFADATQYGTLLVDSSEYFPBEIAKLRDVNGLSLVI FSDWNTSVMRKYKEDR
NTRQMMPTGTGANIPALNELSVNMGSDGLYEGEFLAHNDMYVAGSGIAKPE
DGVNLTQTFKDOGLEYLKOETAVVENVPILGLYQIIPSEGGGRIVLYGDSNCIDDSHRO
KOCFTLLDALLOYTSYGYTPPSLSHGNQRPSSGGLAPPRMBSNHLHRSKYLEA
HIGDPEKRLPACPHLSMAKPOPLNBTASNLKQKLSIDLKVLNFPNSNPQV
RLSPSGSGAMDIPGGIMPGRYNOEVGQTI PVAFGLGAVLAFPVQIISKAKSRPKR
RRPAPRQLQAOHPARPSY"
polyA_signal
4180..4185
/feature="putative"
polyA_site
4198
/feature="putative"
ORIGIN
Alignment Scores:
Pred. No.: 8, 11e-15 Length: 4198
Score: 244.00 Matches: 141
Percent Similarity: 38.7% Conservative: 86
Best Local Similarity: 24.1% Mismatches: 234
Query Match: 7.5% Indels: 126
Gaps: 26
US-10-784-870-4 (1-639) x AK029048 (1-4198)
QY 87 Hs11eVala1GlnPhehanclyPro11leuGluGluThr1yGln1yLeuGluThrThr 106
Db 582 TTTATTTGGCTTTTCAACCGATACCTTCACGCCCAAGCTAGAAACTATTATTTCACAGT 641
QY 107 G1yAla1yV-----11leu 111
Db 642 GCTCTGAAGAAGCAGTGAAGTGAAGAACTGGAATTAATCTCGGAACAACCCATCCAGT 701
QY 112 AepTy11lePProAepTy11leVal1Glu1yGlu1yAepVal1Gln1ySer1yVal 131
Db 702 GACTACCCAGTGAATTTGAG---GTGATTCAAGATTAAGAAGAGACAGAGCGGGGCTG 758
QY 132 ArgSer11leGln1yVal1GlnSerVal1GluProTy11leuProTy11y11leAepPro 151
Db 759 CTCACACTTGAAGATCACCCCAACATCAACGCGGTGACACCCACGCGAAAGCTTTTGGT 818
QY 152 Gln1yPheThr1yGly1yAlaSerThr1yVal1yVal1yAla1yLeu1yAepThr1ySer1n 171
Db 819 TCCCTCAAGTTTCTGAATTCACACCCCATGTCGCC---TGTAAATGAACCCGCGTGG 872
QY 172 AenAen1yGlu1yVal1Gln1yArg1y11leGlu1y11leAlaGln1yTy1yVal1yAlaSer1n 191
Db 873 AGCCAGAAAGTGCGACATCACTCCCGTGAAGAGCAAGCAAGTCTCTCCCTGGGCTCTGGA 932
QY 192 AepVal1y1yTy11leThr1yAla1ySerProGluTy1yTy1yVal1yMetAenAepVal1yAla1yGly 211
Db 933 TTCTGGCATGCAACAGAAACATTCAGATCGGCGATTCGTGAGAGCAATTCCTCGCAG 992
```

```
QY 212 11e-----Val1yVal1AepVal1AlaGlnSerSerTy1yGly1yLeuTy1yGln1y 228
Db 993 GTGCCCGACACACTCCAGCGACAGATGTGCTGG---CAGATGGCATACAGTGTCTAAT 1049
QY 229 Gln1yVal1AlaVal1AaPThrGly1yLeuAepThrGly1yArgAenAepSerSerThs 248
Db 1050 GTCAAGTGTCTGTTTGGATTATCTGGCTC-----AGTACAGACAT 1091
QY 249 Gln1yAepAepGly1yVal1yThr1yAla1yLeuGly1yArgThr1yAenAen1yAen 268
Db 1092 CCGCATTTTAAGAAAT-----GTGAAGAGAGAACCAACATCGACCAAT 1133
QY 269 -----AepTh1yAenGly1y1yGly1yTh1yVal1yAlaGly1ySerVal1yLeuGly 283
Db 1134 GAGCGGACCCCTGATGATGAGGCGGTAGCCATGCGACATTCCTCAGGTCAGG---GTGATTC 1190
QY 284 AenGly1yAen1yGly1yMetA1yProGln1yAen1y---Val1yPheGlnSer11e 302
Db 1191 AGCATGAGGAGTGCACAGATTTCTCCAGATGCGAGCTGCACATCTTCAGGCTCTTT 1250
QY 303 MetAepSerSerGly1yLeuGly1yLeuProSerAen1yLeuGlnThr1yLeuPheSerGln 322
Db 1251 ACCAACAATTCAG-----GTGCTTACACATCTTGCTTGAT 1289
QY 343 Ty1yThr1yAepSerArgAenVal1yAepApy1y-----Val1yArg1yAenAepMetThr 360
Db 1350 GACTTCATGATCATCCCTTTGTTGACAGAGTGGGAAATTAACAGCTTAACATTAATT 1409
QY 361 11leuPheA1yA1yGly1yAenGly1yArgProAenGly1yThr1y1ySer1yA1yProGly1yThr 380
Db 1410 ATGTTTCTGCTATGTCAGATGATGACCTCTGATGCACTGCAATTAACCTGCTGAT 1469
QY 381 Al1yAenAen1y1yThr1yVal1yGly1yAlaThrGln1yAen1yA1yProSerPheGly1ySerTy1y 400
Db 1470 CAGATGAGTGTGATGAGTGGGCTGAC-----TTT 1505
QY 401 Al1yAenAen1yAen1yVal1yAlaGlnPheSerSerArg1y----- 414
Db 1506 GAAGATTAATC-----CTGCGCTTTCTTCGAGGGAAGATCACTCGGGAATTA 1556
QY 415 ProThr1yAepGly1yArg1y1yLeuPheApy1yMetA1yProGly1yThr1yTy1y1yLeuSer 434
Db 1557 CCAGAGGCTATGTCGTGTGACCTGACATGCTGC-----ACCTATGTCGTGCA 1607
QY 435 Al1yAepSerSer1yA1yA1yProAepSerSerPheThr1yAen1yAaPSerTy1yAla 454
Db 1608 GTGCGGGGTTCCGGTGTGAAGAGGGGCTGC-----CGT 1640
QY 455 Ty1yMetGly1yThr1ySerMetA1yThrPro1yVal1yAlaGlyAenVal1yAlaGln1yA1yArg 474
Db 1641 GCATCTCAAGGACCAAGTGTGCTTCCCACTGTCCTGGGCGCTCACTTGTTAATA 1700
QY 475 Gln1yPheVal1yAenA1yGly1y1yThrPro1ySerProSer1yLeu1yVal1yAla1yLeu 494
Db 1701 AGCACAAGTACAGAGCGGAGCTGTG-----AATCTCGACATGTAAGCAAGCTTTG 1754
QY 495 11le1yGly1yA1yAaPVal1yGly1yLeuGly1yTy1yProAenGly1yAenGln1yTy1yGly1yArg 514
Db 1755 ATAGCGTCAAGCCG-GAGACTCTCTGGGGGTCAACATTTGAGCAAGGCTATGG----- 1807
QY 515 Val1yThr1yAep1ySer1yAenVal1yAlaTy1yVal1yAenGly1ySer1yA1yAenSerThr 534
Db 1808 -----CAATTTGATTCGTGCGAGCTTATCAAT 1837
QY 535 SerGln1yA1yA1yThr1yThrPheThr1yAla1yGly1yPhe1yLeu1yV----- 551
Db 1838 CCTCAGCAG-----CTATTAACCCGACAGCAAGCTGAGTCTTACCTAAT 1882
```

QY 552 -----11eserleu-Va1Trp-----SerAspAlaProAlaSerThrThrAl 565
Db 1883 CGACCTGACGATGCTCCCTACATGTCGCCCTACTGCTCCAGCCTATCTACTATGAGG 1942
QY 565 aserValThrleuValaenAspleuAspleuVal11erhAlaProAnGlyThraGly 585
Db 1943 AATGCCAACAATCGTTAT-----GTCCACATCTCTCAATGCGATGGCGGT 1987
QY 585 rValGlyaenAspPheSerAlaProPheAspAenTrp-----AspAl 600
Db 1988 CACAGAAAGAAATTTGGATTAAGCT-----GAGTGGCGAACCTTATTACACAGAA 2038
QY 600 yArGaenAenValGluAenValPhe11eAsenSer-----ProGlnSerGlyTh 616
Db 2039 TGAAGACACATGAAATGCGCTTCTCTACTCTCTCACTGTTGGCCCTGTGAGTTA 2098
QY 616 rTyThr11eGluVal 621
Db 2099 CCTTGCCATCTCCATT 2114

RESULT 5
DQ050980 3159 bp DNA linear GSS 02-JUN-2005
LOCUS Homo sapiens MBTPS1 gene, VIRUTAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION DQ050980
VERSION DQ050980.1 GI:66904179
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3159)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Piedel-Alon,A., Tanenbaum,D.M., Civeello,D.,
White,T.J., Sinsky,J.J., Adams,M.D. and Cargill,M.
A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
(er) PLoS Biol. 3 (6), E170 (2005)
15869325
JOURNAL 2 (bases 1 to 3159)
PUBMED
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,
Hubisz,M.J., Piedel-Alon,A., Tanenbaum,D.M., Civeello,D.,
White,T.J., Sinsky,J.J., Adams,M.D. and Cargill,M.
Direct Submision
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.

FEATURES
source 1. 3159
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
<1..>3159
/gene="MBTPS1"
/locus_tag="HC19002"

ORIGIN
Alignment Scores:
Pred. No.: 6.6e-15 Length: 3159
Score: 243.00 Matches: 140
Percent Similarity: 38.6% Conservative: 89
Best Local Similarity: 23.6% Mismatches: 224
Query Match: 7.4% Indels: 140
DB: 11 Gaps: 28

US-10-784-870-4 (1-639) x DQ050980 (1-3159)
QY 87 H1s11eValGlnPheAnGlyPro11eLeuGluGluThrLysGlnLysLeuGluThrThr 106

Db 166 TAAATGTGCTTCAATGATATCTTACAGCCAAAGCTAAGAAATCATTTATTTCAGT 225
QY 107 GlyAlaLys-----11eLeu 111
Db 226 GCCCTGAAGACGATGAAGATGACAAATGGAGAAATTATCTCGAAACAATCCATCCAGT 285
QY 112 AspTyThr11eProAspTyThrAlaTyThr11eValGluTyGluGlyAspValGlnSerLysVal 131
Db 286 GACTACCTTAAGTATTTTGAAG--GTGAATCAATTAAGAAAGAAAGAAAGCGGGCTG 342
QY 132 ArgSer11eGluThiValGluSerValGluProTyLysProLysTyThr11eAspPro 151
Db 343 CTAACTTAAGATATATCCAAACATCAAAACGGGTACACGCCCAAGAAAGCTTTCTGT 402
QY 152 GlnLeuPheThrLysGlyAlaSerThrLysValLysAlaLeuAlaLeuAspThrLysGln 171
Db 403 TCCCTC-----AAGTATCTGAATCTGACCCCAAGTACCTGCAATGAACCCGGTGG 456
QY 172 AsnAsnLysGluValGlnLeuArgGly11eGluGlu11eAlaGlnTyThrValAlaSerAsn 191
Db 457 AGCCAGAAAGTGCATATCATCAGTCCCTGGAAAGCCAGCCTCTCCCTGGGCTCTGGC 516
QY 192 AspValThiGlyThr11eThrAlaLysProGluTyThrLysValMetAsnAspValAlaArgGly 211
Db 517 TTCTGGCATCTCAAGGAAAGCATTTGAGACAGACAGCGCTGTGAGACCATCCGCCCGAC 576
QY 212 11e-----ValLysAlaAspValAlaGlnSerSerTyGlyLeuTyGlyGlnGly 228
Db 577 GTTGCCCAAGACATCGACAGGACAGATGTGCTGG--CAGATGGAGATACAGGTGCTAAT 633
QY 229 Gln11eValAlaValAlaAspThrGlyLeuAspThrGlyAlaArgAsnAspSerMetHis 248
Db 634 GTAAAGTGTGCTTTTTCACATGCGCTG-----AGCGAAGACAT 675
QY 249 GluAlaPheArgGlyLysThr11eThrAlaLeuTyThrAlaLeuGlyAlaArgThrAsnAlaAsn 268
Db 676 CCCCACTTAATAAT-----GTAAAGAGAGAACCAATCGAGACCAAC 717
QY 269 -----AspThrAnGlyThiGlyThiThiLeuAlaLysSerValLeuGly 283
Db 718 GAGCGAAGCTGAGACATGAGTGGGCCATGACATTCGTGGCAGGT--GTGATAGGC 774
QY 284 AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeu--ValPheGlnSer11e 302
Db 775 AGCATGAGGAGATGCCAAGATTTGCTCAATGACGAATCTCACTTTCAGGGCTCTT 834
QY 303 MetAspSerSerGlyLysLeuGlyLysLeuProSerAsnLeuGlnThrLeuPheSerGln 322
Db 835 ACCAATATACAG-----GTATCTTACATCTTGCTTTTGGAC 873
QY 323 AlaPheSerAlaGlyAlaArg11eThiThiAsnSerTyGlyAlaAlaValaenGlyAla 342
Db 874 GCCTTCAACTATGACATTTTAAAGAAATGACAGTGTAACTCAGCATCGCGCGCG 933
QY 343 TyThrThrAspSerArgAsnValaAspAspTyThr-----ValArgLysAsnAspMetThr 360
Db 934 GACTTCATGATCATCCGTTGTTTGAACAAGGTGGGAATTTAACAGCTTAACATGTAATC 993
QY 361 11eLeuPheAlaLysGlyAsnGlyLysArgProAsnGlyGlyThi11eSerAlaProGlyThr 380
Db 994 ATGGTTTCTGCTAATTTGGCAATACGAGACCTCTTATGCGACTGTGAATACCTGTCTAT 1053
QY 381 AlaLysAsnAla11eThrValaGlyAlaThrGluAsnLeuArgProSerPheGlySerTy 400
Db 1054 CAATGAGATGATGAGATGAGTACGCGGCATTTAC-----TTT 1089
QY 401 AlaAspAsn11eAsnHisVala11eGlnPheSerSerArgGly----- 414
Db 1090 GAAGATTAACATC-----GCCGCTTTTTCAGAGGGAATGACTACTGGAGACTA 1140
QY 415 ProThrLysAspGlyAlaGlyLysProAspValMetAlaProGlyThiTyThr11eLeuSer 434

Db 1141 CCAAGAGCTACAGCTGCATGAACCTGACATTGTC-----ACCTATGCTGCTGC 1191
Qy 435 A1AATGSerSerLeuAlaProApeSerSerPheThrAlaAsnHisApeSerIleTyrAla 454
Db 1192 GTGCGGGGTTCTGGCCTGAAAGGGGGTGC-----CGG 1224
Qy 455 TymeG1yG1YThSerMetAlaThrProIleValAlaG1YAsnValAlaG1nLeuAtrg 474
Db 1225 GCCCTCTCAGGAGCAAGCTGCTTCTCCAGTGGTTCAGAGTCTCTCACCCTGTTAGTG 1284
Qy 475 GluHisApeValIleApeAsnArgIleThrProIleProSerIleuLeuValAlaLeu 494
Db 1285 AGCACAGTCCAGAAAGGCTGAGCTGTG-----AATCCCGCATATGAAGCAGGCCCTG 1338
Qy 495 I1eAlaG1YAlaAlaApeValAlaG1YLeuG1YThrProAsnG1YAsn-----G1nG1Y 511
Db 1339 ATCGGCTCAGCCCGAGGCTC-----CCCGGGGTCAACATGTTTGAGCAAGGC 1386
Qy 512 TrpG1YArgValIleLeu-----ApeIle 519
Db 1387 CACGGCAAGCTGCATCTGCTTAGAGCTATCAGATCTCAAGCAAGCTCAAGCAGCA 1446
Qy 520 SerLeuAsnValAlaTyrValAsnG1nSerSerAla-----LeuSerThrSer 535
Db 1447 AGTTTGAAGCCCAAGCTACATAGATCTGAGTGTCCCTACATGTGGCCCTACTGCTCC 1506
Qy 536 GluIleValAlaThrTyrThrPheThrAlaThrAlaG1YIleProLeuIleIleSerIleVal 555
Db 1507 CAGCCCATCTACTA-----GGAGAAATGCCG----- 1533
Qy 556 TrpSerApeAlaProAlaSerThrThrAlaSerValIleLeuValAsnApeLeuApeLeu 575
Db 1534 -----ACAGTTGTATATGTCACATCTCTCAAGCGCATGGAGATC 1572
Qy 576 -----ValIleThrAlaProAsnG1YThrArgTyrValG1YAsnApePheSerAla 592
Db 1573 ACAGAGAAATGTGATATAGCTGACCTGAGCCCTATTG----- 1614
Qy 593 ProPheApeAsnAsnTrpApeG1YArgAsnAsnValG1uAsnValPheIleAsnSer--- 611
Db 1615 -----CCACAGAACGAGAAC-----AACATGTAAAGTTGCCCTCTCTACTCTCTCG 1659
Qy 612 -----ProG1nSerG1YThrTyrThrIleG1uVal 621
Db 1660 GTCTTATGAGCTGTGTCGGGCTACCTGGGCATCTCCATT 1698
RESULT 7
LOCUS DQ050981 3159 bp DNA linear GSS 02-JUN-2005
DEFINITION Pan troglodytes MBTPS1 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ050981
VERSION DQ050981.1 GI:66904180
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Bukaryaia, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
REFERENCE
AUTHORS 1 (bases 1 to 3159)
Meisen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeille, D.,
White, T.J., Sniinsky, J.J., Adams, M.D. and Cargill, M.
A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
(et) Plos Biol. 3 (6), E170 (2005)
JOURNAL
PUBMED 15869325
REFERENCE 2 (bases 1 to 3159)
AUTHORS Meisen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeille, D.,
White, T.J., Sniinsky, J.J., Adams, M.D. and Cargill, M.
Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source location/Qualifiers
1..3159
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>3159
/gene="MBTPS1"
/locus_tag="HC19002"
ORIGIN
Alignment Scores:
Pred. No.: 6.6e-15 Length: 3159
Score: 243.00 Matches: 140
Percent Similarity: 38.6% Conservative: 89
Best Local Similarity: 23.6% Mismatches: 224
Query Match: 7.4% Indels: 140
DB: 11 Gaps: 28
US-10-784-870-4 (1-639) x DQ050981 (1-3159)
Qy 87 HisIleValG1nPheAsnG1YProIleLeuG1nG1uThrIleG1nIleLeuG1uThr 106
Db 166 TATATTGGCTTTCATATGATCTTATACAGCCAAAGCTAGAAATTCATTATTCAAGT 225
Qy 107 GluAlaIle-----IleLeu 111
Db 226 GCCCTGAAGACAGTGAAGTGAACAATTGAGAAATTATCTCGAAACATCATCCAGT 285
Qy 112 AspTyrIleProApeTyrAlaTyrIleValG1uTyrG1uIleApeValG1nSerIleVal 131
Db 286 GACTAACCTACTGATTTTGTAG--GTGATTCAGATTAAGAAACGAAACGCGGGCTG 342
Qy 132 ArgSerIleGluHisValG1uSerValG1uProTyrIleuProIleuPyrIleApePro 151
Db 343 CTAACTGAAAGATCATCCAAACATCAAGGGGTACGCCCCCAAGAAAGTCTTTCGT 402
Qy 152 GluLeuPheThrIleG1YIleAsnSerThrLeuValIleValAlaLeuApeThrIleG1n 171
Db 403 TCCCTC-----AGTATGCTGAATCTGACCCCAAGTACCTGCTCAATGAACCGGTG 456
Qy 172 AsnAsnIleG1uValG1nLeuArgG1YIleG1uIleAlaG1nTyrValAlaSerAsn 191
Db 457 AGCCAGAAAGTGCATATCATCACTCCCTCGAAGACCAAGCTCTCTGGGCTCTGC 516
Qy 192 AspValHisTyrIleThrAlaIleProG1uTyrIleValMetAsnApeValAlaArgG1Y 211
Db 517 TTTCGCAATGCTACGGAGAGCATTCAGACAGACGCTGCTGAGACCATCCGCGCAG 576
Qy 212 Ile-----ValIleAlaApeValAlaG1nSerSerTyrG1YLeuTyrG1nG1Y 228
Db 577 GTTCCCAAGACATCGAGGAGATGTCTCTG--CAGATGGATATACAGGTCTAAT 633
Qy 229 GluIleValAlaValAlaApeThrG1YLeuApeThrIleArgAsnApeSerMetHis 248
Db 634 GTAAGATGTCTGTTTGTGACACGTGGCTG-----AGCGAGAACAT 675
Qy 249 GluHisApeAsnG1YIleThrAlaIleuTyrAlaLeuG1YArgThrAsnApeValAsn 268
Db 676 CCCCACTTCAAAAT-----GTGAAGAGAACCAACTGACCAAC 717
Qy 269 -----AspThrAsnG1YIleG1YThrIleValAlaG1YSerValIleG1Y 283
Db 718 GAGCGAAGCTGATGATGGGTGGCCATGAGCATTCGTCGAGGT---GTATAGCC 774
Qy 284 AsnG1YAlaThrAsnIleG1YMetAlaProG1nAlaAsnLeu---ValPheG1nSerIle 302
Db 775 AGCATGAGGAGTGCAGAGATTTGCTCCAGATCAACATTCACATTTCCAGGGCTTT 834
Qy 303 MetApeSerSerG1YIleuG1YIleuProSerAsnLeuG1nThrIleuPheSerG1n 322

```

Db      835 ACCAATATATGAG-----GTATCTACACATCTGTTTGGAC 873
Qy      323 AAlPhSeRAlaGlYAlaRgIleHlStPhSeRlTrpGlyAlaValaLaenglyAla 342
Db      874 GCCTTCACCTATCCATTTTAAAGAGATTCAGTGTAAACCTCAGCATCGCGCGCCG 933
Qy      343 TTrThTrAspSerArGanValaAspArTyR-----ValaRgYsAsnAspMetTr 360
Db      934 GACTCATGATCATCCCTTTGTTCACAGCTGTGGAAATTAACAGTAAACATGATATC 993
Qy      361 lIeLeuPheAlaAlaGlYanGlYuarProanGlYglYThrIleSeRAlaProGlYlTr 380
Db      994 ATGTTTCCTATTCAGTACAGGACCTTTATGSCACTGTATATACCTTCGAT 1053
Qy      381 AlAlaYasAlaIleThrValGlYAlaThrGlYAsnLeuArgProSeRPhelGlySerTyR 400
Db      1054 CAATGATGTGATTGGATGAGCGCATTCAGC-----TTT 1089
Qy      401 AlAlaAsnIleAsnHlStValAlaInPhSeSerArGly----- 414
Db      1090 GAAGATTAATC-----GCCCGCTTTCTTCAAGGGAAATGACTACCTGGAGCTA 1140
Qy      415 ProThTrLyAspGlYArGlYleYsProAspValMetAlaProGlYlTrYlIeUser 434
Db      1141 CCAAGAGGCTACGCTGCATGAACCTGACATGTC-----ACCTATGGTGTCTGGC 1191
Qy      435 AlAlaRSeSerLeuAlaProAspSerSerPheTrpAlaAsnHlStAspSerLySlyAla 454
Db      1192 GTCCGGGGTCTTCGCGTGAAGGGGGTGC-----CGG 1224
Qy      455 TyrMetGlYlYThrSeMetAlaThrProIleValAlaGlYAsnValAlaInLeuArg 474
Db      1225 GCCCTTCAGGAGCCAGGTGTCTCTCCAGTGTGACAGTCTGTCACTTAATTAAGTG 1284
Qy      475 GlUHLPhEValIYsAsnArGlYlIeThrProLyProSeRLeuLeuYsAlaIleu 494
Db      1285 AGCAGAGTCAGAGAGGTGAGCTGGTG-----ATCTGCCAGTATGAAGAGGCCCTG 1338
Qy      495 lIeAlaGlYAlaAlaAspValGlYleuGlYlYrProanGlYAsn-----Gingly 511
Db      1339 ATCGCTCAGCGCGGAGGCTC-----CCGGGGGTCAACATGTTTGAGCAAGGC 1386
Qy      512 TrpGlYArGValThreU-----AaplyS 519
Db      1387 CACGGCAAGCTCATGCTCTCAGAGCCTATCAGATCTCAACAGCTAACAGCCAGGCA 1446
Qy      520 SerLeuAsnValAlaIYrValaLaenglYSeRAla-----LeuSerThrSer 535
Db      1447 AGTTTAAAGCCCGACATCATGATCTGACGTAGTCCCTACATGAGCCCTTACTGCTCC 1506
Qy      536 GlNLYsAlaIleThrTrhPheThrAlaThrAlaGlYlYsPheLeuYsIleSeRleuVal 555
Db      1507 CAGCCCATCTACTAT-----GAGGAATGCCG----- 1533
Qy      556 TrpSerAspAlaProAlaSerThrThrAlaSerValThreUValaAsnPhleuAspLeu 575
Db      1534 -----ACAGTGTTAATGTCAATCTCTCAACGCGCAGGAGGAGTC 1572
Qy      576 -----ValIleThrAlaProanGlYlYrArGTYrValGlYlYsAsnPhSeRAla 592
Db      1573 AACAAGAAATTTAATAAGCTTGAAGCTGACGCCCTATTTG----- 1614
Qy      593 ProPheAspAsnAsnTrpAspGlYArGAsnAsnValGlYAsnValPheIleAsnSer--- 611
Db      1615 -----CCACAGAACGAGAGC-----AACATTGAAGTTGCTTCTCTACTCTCTCG 1659
Qy      612 -----ProGlnSerGlYlYrThrTrhIleGlYVal 621
Db      1660 GTCTTATGAGCTTGTGGGCTACCTGAGCATCTCATTT 1698

```

```

LOCUS      CM807764          679 bp    DNA          linear    GSS 20-JAN-2005
DEFINITION pIne_89392 Phytophthora infestans T30-4 Phytophthora infestans
ACCESSION  Genomic, genomic survey sequence.
VERSION    CM807764
KEYWORDS   CM807764.1 GI:57982181
SOURCE     Phytophthora infestans (potato late blight agent)
ORGANISM   Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
            Phytophthora.
REFERENCE  1 (bases 1 to 679)
AUTHORS   Randall,T.A., Dwyer,R.A., Huitema,E., Beyer,K., Cvitanich,C.,
            Kellar,H., Fong,A.M., Gates,K., Roberts,S., Yatzkan,B., Gaffney,T.,
            Law,M., Teeter,A., Torto-Alalibo,T., Zhang,M., Zheng,L., Mueller,B.,
            Windas,J., Binder,A., Birch,P., Gisl,U., Govere,F., Gow,N.A.,
            Mauch,F., van West,P., Maugh,M.E., Yu,J., Bolter,T., Kamoun,S.,
            Lam,S.T. and Judelson,H.S.
            Large-scale gene discovery in the oomycete Phytophthora infestans
            reveals likely components of phytopathogenicity shared with true
            fungi. Plant-Microbe Interact. 18 (3), 229-243 (2005)
JOURNAL    Mol. Plant-Microbe Interact. 18 (3), 229-243 (2005)
PUBMED     15782637
COMMENT    Contact: Judelson HS
            Department of Plant Pathology
            University of California
            PUEP, Riverside, CA 92521, USA
            Tel: 9518274199
            Fax: 9518274294
            Email: howard.judelson@ucr.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
            source          1..679
                        /organism="Phytophthora infestans"
                        /mol_type="genomic DNA"
                        /isolate="T30-4"
                        /db_xref="taxon:4787"
                        /sex="M1"
                        /issue_type="total mycelium"
                        /clone_lib="Phytophthora infestans T30-4"
                        /note="Vector: pUC18"
ORIGIN
Alignment Scores:
Pred. No.:      6,16e-16      Length:      679
Score:          242.50      Matches:      76
Percent Similarity: 46.7%      Conservative: 23
Best Local Similarity: 35.8%      Mismatches:  70
Query Match:    7.4%      Indels:      43
DB:             10      Gaps:      8
US-10-784-870-4 (1-639) x CM807764 (1-679)
Qy      273 HlAgLYThrHlValAlaGlYSeRValLeuGlYAsnGlYAlaThrAsnLyglYMetAla 292
Db      3 CACGGTACACAGTATACGGGAATTAATGTGGGCCAA-----TTGGGCTTCGAGAGCGCA 56
Qy      293 ProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlYlYleuGlYlYleu 312
Db      57 CCGGAGGCTACCTGATGATGATGAAGGATGTCGTCGCTCAAGGTTGCTATGGG----- 110
Qy      313 ProSerAsnLeuGlnThreUpe----- 320
Db      111 ---TCGATTGCTAGCATGTTTCCAGTTCATGTTGTGCTTACGACACCCAGTGGACA 167
Qy      321 SerGlnAlaPheSerAlaGlYAlaRgIleHlStPhSeRlTrpGlyAlaValaLaen 340
Db      166 TCAAAAGACTGCTCCAAAGACCGAATATTGTTAACAACAGTTGGGT-----GGC 218
Qy      341 GlYAlaTYrThTrAspSerArGanValaAspArTYrValArGlyAsnAspMetTr 360
Db      219 GGTCAAGGTCTCACGATGTTTATGAGGGTTATCAACGCTTGGCAGAGTGCAGGATCATTT 278
Qy      361 lIeLeuPheAlaAlaGlYanGlYuarProanGlYglYThrIleSeRAlaProGlYlTr 380

```

[illegible]

Db 82 TTCTTTGAGATTCAAAGTATCCATACCATTTAATCAAGTGAATGAAAAATCTAGAAAA 141

Qy 254 IysIIeThAlaLeuThrAlaLeuGlyA-gThrAsnAsnAlaAsnAspThrAsnGlyHis 273

Db 142 GTTGTAACT-----TATATTAATCTTACCATGACAAATGAAGATTATGTAAATGGTCAAT 192

Qy 274 GLyTHrHisValAlAGlySerValLeuGlyAsnGly-----Ala 286

Db 193 GGTACACATGTTTGCGCTCTCCAGCAGGTACTCCAGAGATTCCTCATGGGCTAATTCA 252

Qy 287 ThrAsnGlyMetAlaProGlnAlaAsnValPheGlnSerIleMetAspSerSer 306

Db 253 TCATTTAAGTGCGTCTTGCAACTGACACCAAGATTCGATTTTAAGATCTT-----TATCT 306

Qy 307 GlyGlyLeuGlyGlyLeu---ProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325

Db 307 GGAAGTCTGCAACCAACCAACCCGAAAGATTACGTCAATGATCAAAATGCAACCATTAATGAT 366

Qy 326 AlaGlyAlaValGlnIleHisThrAsnSerTrpGlyAla-----AlaValAsnGly 341

Db 367 GCAAGTGCAGAAAGTACATGATGATCTTGCGGGTCTGTATCTTGCAACAGTTATTAATGAT 426

Qy 342 AlaYrThrThrAspSerArgAsnValAspAspYrValArgLys---AsnAspMetThr 360

Db 427 GGTATTTCCGATGATGCTCGTGATGATTAATGATGATCTCTTAAGATGACCAAGATTTCTT 486

Qy 361 ILeuLeuPheAlaAlaGlyAsnGlnArgProAsnGlyGlyThrIleSerAlaProGlyThr 380

Db 487 ATACTAAAGAGCTGCTGTATTAATACAG---CTAATTGCATCTTTATTGACTCAAGCA 543

Qy 381 AlaYrAsnAlaIleThrValGlyAlaThrGlnLeuAspArgProSerPheGlySer--- 399

Db 544 GCTAATAATGCAATTACATGATGGTGCTGAGCAAAAGCCTCATGTAAATTATGTGTCAAT 603

Qy 400 -----TyrAlaAspAsnIleAsn 405

Db 604 GCATTGGAATATTAATGATTTCTCAAGTAAATGCTAAT 639

RESULT 10

LOCUS BJ393752 594 bp mRNA linear EST 08-MAR-2000

DEFINITION BJ393752 Dictyostelium discoidium cDNA library, SF Dictyostelium discoidium cDNA clone dds32b16 5', mRNA sequence.

ACCESSION BJ393752

VERSION BJ393752.1 GI:19304838

KEYWORDS EST.

SOURCE Dictyostelium discoidium

ORGANISM Dictyostelium discoidium

REFERENCE Bukaryota, Mycetozoa; Dictyostellida; Dictyostelium.

AUTHORS I. (bases 1 to 594)

TITLE Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.

JOURNAL Full length cDNA of Dictyostelium discoidium at the s1ug stage unpublished (2002)

COMMENT Contact: Tadasu Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tehin@genes.nig.ac.jp.

FEATURES

source

1..594

/organism="Dictyostelium discoidium"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="dds32b16"

/sex="mat A"

/dev_stage="S1ug stage"

/clone_lib="Dictyostelium discoidium cDNA library, SF"

Pred. No.: 1,41e-13 Length: 594
 Score: 220.50 Matches: 54
 Percent Similarity: 51.4% Conservative: 40
 Best Local Similarity: 29.5% Mismatches: 64
 Query Match: 6.7% Indels: 25
 DB: 3 Gaps: 7

US-10-784-870-4 (1-639) x BJ393752 (1-594)

QY 254 LysileThraAsenValAlaLeuGlyAArgThraAsnAlaAsnAphThraAsnGlyHis 273
 DB 21 AAGGTGTAACTTATATTACATCAATCAACGACGACATGATGATTAAGATGATGCTAC 80
 QY 274 GlyThraHisValAlaGlySerValLeuGly-----AanglyAla 286
 DB 81 GGTAACATATTTGTGGTCTGCAGCAGGATCTCCAGAGATTTCTCAATTAATTTATTT 140
 QY 287 ThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSer 306
 DB 141 TCATTAGTGTCTTCCACTGATGCAAGATTCATCTTTGATTTGGCAAGTGTCTCA 200
 QY 307 GlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThraPheSerGlnAlaPheSerAla 326
 DB 201 TCAAGTTTGAACCT---CCATCGATTTGAACAATATATATACACCATATATGACGCA 257
 QY 327 GlyAlaArgGlyLeuThraAsnSerTGGlyAla-----AlaValaanglyAla 342
 DB 258 GGTCAGAGGTGATGATGATTTCTTGGGCTTCTGATACGATGAGGGGTATACAGGTATG 317
 QY 343 TTTThraAspSerArgAsnValAlaAspPyrValArgLysAsn---AspMetThrIle 361
 DB 318 TATTCATCAGACACCTCTCATATGATGATTTCTTTTCTCATCTCATCCAGATTTTCATCAT 377
 QY 362 LeuPheAlaAlaGlyAnglyLeuArgProAsnGlyGlyThraIleSerAlaProGlyThraAla 381
 DB 378 CTAGAGCTGCTGTGAACAACGACGATACCTATCATCTACT---CAATCAGCTGCA 434
 QY 382 LysAsnAlaIleThraValGlyAlaThraGlnAsnLeuArg----- 394
 DB 435 AAGAAATGTTATTAACCTGTCTCATCAACAATTCATGAAATTAATTAATTAATGATGAT 494
 QY 395 ProSerPheGlySerTyrAlaAsnAlaAsn-----HisValAlaGlnPhe 410
 DB 495 CCAATATATATTAATTTATCATCATCTCTGATTAATCAACAGATTAATATGATTTTC 554
 QY 411 SerSerArg 413
 DB 555 GATACGACA 563

RESULT 11
 LOCUS CDS0A30A 1695 bp mRNA linear HTC 06-FEB-2004
 DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 Arabidopsis thaliana (thale cress).
 BX827771 GI:42460584
 VERSION
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1695)
 REFERENCE
 AUTHORS Caselli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
 Menard,M., Cruaud,C., Queller,F., Scarpelli,C., Schachter,V.,
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
 TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1695)
 AUTHORS Genoscope.

TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Caselli
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
 source
 Location/Qualifiers
 1..1695
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSLRPH1ZP09"
 /rname_type="Hormone Treated Callus"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"
 1..1695
 /gene="At4g00230"

ORIGIN
 Alignment Scores:
 Pred. No.: 1e-12 Length: 1695
 Score: 219.50 Matches: 128
 Percent Similarity: 33.7% Conservative: 55
 Best Local Similarity: 23.6% Mismatches: 153
 Query Match: 6.7% Indels: 207
 DB: 4 Gaps: 25

US-10-784-870-4 (1-639) x CDS0A30A (1-1695)

QY 262 GlyArgThraAsnAlaAsnAphThraAsnGlyHisGlyThraHisValAlaGlySerVal 281
 DB 1 GGCGAAGTCCGATCACCGATTCATATCGACGACATGGAACGACACGTCATCACCGTA 60
 QY 282 LeuGly-----AsnGlyAlaThraAsnLys 289
 DB 61 GGCGGCTTTTATGTCGCAACGCGAGTCTATGCGATGACAAACGCG---ACCGCCCG 117
 QY 290 GlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyLeu 309
 DB 118 GGCGCGGTTCCGTCGCGGAGGATGATCAAGATTGTTGGCGGAGATCCGCGTGT 177
 QY 310 GlyGlyLeuProSerAsnLeuGlnThraPheSerGlnAlaPheSerAlaGlyAlaArg 329
 DB 178 GCTGACATG-----GACATTACTCGCCGATTCGAACGCGATTCACACGCGGTGA 231
 QY 330 IleHisThraAsnSerTyrGlyAlaAlaValaanglyAlaTyrThraAspSerArgAsn 349
 DB 232 ATTATCTCATCTATTCGCGCGTCCATC---GCCGATTACTCTTCCGATTCGATATCC 288
 QY 350 ValAspAspTyr-----ValArgLysAsnAphMetThrIleLeuPheAlaAlaGlyAsn 367
 DB 289 GTCCGGCTCTTTCACGCGATGAGGAAGATATCTCACGCGTGGCG---TCCGCGCGTAC 345
 QY 368 GlyArgProSerAsnGlyGlyThra----- 374
 DB 346 GACGGGCTAGTCTCAAGAACTGTAACGAACATGACCGGTGATTTGACGGTGTCTGCA 405
 QY 374 ----- 374
 DB 406 AGTGAATGATCGAGCTTCAAGACAAATATGATCTCGGCAACGCAATCTTCTCT 465
 QY 375 -----IleSerAlaPro 378

```

Db      466 GGGATGGGAATAGCATGTTTAGTCCAAAAGCCAAATCGATCCCTGTAGTGTGTT 525
Qy      379 G|YThrAlaIyAsnAla-----:::|
Db      526 GATGCTGCTAAGAACAGACATAGTACTTGGCTAGGTATGTTCTCTCATTTCTTG 585
Qy      385 -----IleThrValGlyAlaThrGlu 391
Db      586 GATCGAAAGAGTGAAGGAAAGGTATGTGTAGATGAGAGAGTGTGTGAG 645
Qy      392 AenIuAaArgProSerPheGly-----SerTyrAlaAsp 402
Db      646 TCTACTATCAAAAGCTATGAGAGTGTGTGTCATCATTTGATATCAATATCTTGAC 705
Qy      403 Aen-----403
Db      706 AATGCTCAATTTTCATGCGACCTGCGACCAAGTTTAATAGTCGCTGGCATATATAC 765
Qy      404 -----IleAsn-----405
Db      766 TACCGATATATCAATCCCAAGATCAGCGTGGCTGTTATTCAGAAACTCGGCAAGTG 825
Qy      406 -----HisValAlaGlnPheSerSerArgGlyProThrIyAsnGlyArg 420
Db      826 ACAATCCCTGCTCCATTTGCTGCTTTTTCATCAAGAGTCC--AATCCGGATCA 882
Qy      421 Ile-----LysProAspValMetAlaProGlyThrTyrIleLeuSerAla-----435
Db      883 ATAACGCTTTTCAAGCTGATATCGCTGACCGGGATGATATATATGCGCGCTTAATCT 942
Qy      436 -----ArgSerSerIleuAlaProAspSerSerPheThrAlaAsnHisAspSer 451
Db      943 CTAAAGAGTAACCTGCTGGTTAGATGTGTACACCCAGTCTC-----TCA 987
Qy      452 LysTyrAlaIyAsnGlyIyThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
Db      988 AATTTACCATCTCTGTTGGACCTCAATGCGCTGCGCTCATGTTCTGTGTAGTGTGG 1047
Qy      472 GlnLeuArgGlnIlePheValIyAsnArgGlyIleThrProIyAsnLeuLeuIyAsn 491
Db      1048 -----TACGTCAGTCTTTTTCATCCGATTTGACCCCTGCTGCGCATCAAA 1092
Qy      492 AlaAlaIleuIleAlaGlyAlaIleAspValGlyLeuGlyTyrProAsnGlyAsnGly 511
Db      1093 TCGGCATCATTAATCTCAGCAAAACCGATTAAGCCGG-----1128
Qy      512 TrpGlyArgValThrIleuAspIySerIleuAsnValAlaTyr-----ValAsn 527
Db      1129 -----AGAGTGAACAAAGACGCAAGATTTGCTTATGAGAGAGCCAAATTAAC 1176
Qy      528 GluSerSerAlaLeuSerThrSerGlnIyAsnAlaThrTyr-----PheThr 543
Db      1177 CCACGACGAGCCCGACAGCTGCTTAGTACACATGACGACATCTCTATGTTCAAGT 1236
Qy      544 AlaThrAlaGlyIyAsn-----ProLeuIyIleSerLeuValTrpSerAspAla 559
Db      1237 TCTTGTGCGGCGAAGGCTACACGCAACGACTTACCTCATTTGG--TGGGACACGCT 1293
Qy      560 ProAlaSerThrThrAlaSerValThrIleuValAsnAspLeuAspLeu--ValIleThr 578
Db      1294 CCGTGAGCT-----GTTCCCTCCATTTGCTCCGTCGAGCGGCGACAG 1332
Qy      579 AlaProAsnGlyThrArgTyrValIyAsnAspPheSerAlaProPheAspAsnAsnTrp 598
Db      1333 ATTCCCTCACTACCCCAATCAACTCACTGATGAGATCCGCCAAACCTCCACATTTG 1392
Qy      599 --AspGlyArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrT 618
Db      1393 CTGTGTTCAGGCGGAGAGTCAACACGTG-----GGAACACGCTGCTCGGTCTACA 1443
Qy      618 hrIleGluValAlaIle-----TyrAsnValProValGlyProGlnAsnAspSer 635

```

```

Db      1444 CCGCCACCGTCCGAGACCCGAAAGAGTGAATACGCTGAGAGCACAGATTGTTCAT 1503
Qy      635 euAla 636
Db      1504 TTTCa 1508

RESULT 12
LOCUS   DR651450/c 962 bp mRNA linear EST 12-JUL-2005
DEFINITION EST1041567 FVN Gibberella moniliformis cDNA clone FVN015, mRNA
sequence.
ACCESSION DR651450
VERSION   DR651450.1 GI:70739926
KEYWORDS EST.
SOURCE   Gibberella moniliformis
ORGANISM Gibberella moniliformis
          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
          Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 962)
          Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
          Utecher, F., Smith, S., Feldblyum, T., Glenn, A.B., Plattner, R.D.,
          Kendra, D.F., Town, C.D. and Whitebaw, C.A.
          Analysis of 87,000 expressed sequence tags reveals alternatively
          spliced introns in multiple genes of the fumonisin gene cluster
          unpublished (2005)
JOURNAL  Contact: Brown, D.W.
COMMENT   USDA/ARS/NCAR
          USDA
          1815 N. University St, Peoria, IL 61604, USA
          Tel: 309 681 6230
          Fax: 309 681 6689
          Email: brown@ncar.usda.gov
          TIGR sequence name: FVN015TV
          Seq primer: GTA ATA CGA CTC ACT ATA GGG C.
          Location/Qualifiers
            1..962
              /organism="Gibberella moniliformis"
              /mol_type="mRNA"
              /strain="m3125"
              /db_xref="taxon:117187"
              /clone="FVN015"
              /tissue_type="mycelia"
              /clone_id="FVN"
              /note="Vector: pBluescript II SK(+) XR; Site 1: EcoRI;
              Site 2: XhoI; anamorph: Fusarium verticillioides. Library
              FVN was obtained from RNA derived from a corn meal medium
              culture of strain M-3125. These cultures were prepared by
              inoculating an autoclaved mixture of 25 g corn meal and 5
              ml distilled water with 5 ml of water containing 5 x 10e7
              conidia. The inoculated medium was mixed thoroughly,
              distributed equally into two 100-mm plastic petri dishes,
              and incubated at room temperature for four or six days."

ORIGIN
Alignment Scores:
Pred. No.: 4,53e-13 Length: 962
Score: 219.00 Matches: 87
Percent Similarity: 44.4% Conservative: 40
Best Local Similarity: 30.4% Mismatches: 109
Query Match: 6.7% Indels: 50
DB: 8 Gaps: 15

US-10-784-870-4 (1-639) x DR651450 (1-962)

Qy      233 ValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArg 252
Db      951 GTGTGACACATGTGTTCGAAC-----ACCATATGAAATTTGAG 910
Qy      253 GlyIyIleThrAlaIleuTyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGly 272
Db      909 GGCCTGCTCAGGCGCTTTATCTGCTTCATGCTGAGTGAACAAGCC--GATAGTGTGGC 853
Qy      273 HisGlyThrHisValAlaGlySerValLeuGlyAsn-----GlyAlaThrAsnIyGly 290

```



```

Db      852 CACGGTACTACGCTCTGAGACATTCGCGGCAAGCACTACGCTGTCTCCAAAGGCC 793
Qy      291 MetAlaProGlnAlaLeuValPheGln--SerIleMetApsSerSerglyLeu 309
Db      792 ACCATC---CAAGCTGTCAAGGTCTTCCAGGCGCTCTTCTAGACCTCATCATCTT 736
Qy      310 GtlyGlyLeuProSerAenLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArg 329
Db      735 GCTGGTTTCAACTGGCTGCCAAGATATCATCTCCAAAGGCCAAGCTGCTCGCTCGTC 676
Qy      330 IleHisThrAenSerTrpGlyAlaAlaValaasnGlyAlaTrpThrApsSerArgasn 349
Db      675 GTGAAC-----ATGCTCTGCGTGGTGTTACTCTGCTTCTTCAACAC 631
Qy      350 ValApsAerGlyValAlaGlyAenApsMetThrIleLeuPheAlaAlaGlyAenGlyAlaArg 369
Db      630 GCTGTCAACTGTCTCTCATCTGATCATCTCTCCATTCGCTGGAAGCATGGT 571
Qy      370 ProasnGlyGlyThrIleSerAlaProGlyThrAlaApsAlaIleThrValAla 389
Db      570 GCCAAGCGCCCAACACTCT---CTGGCTGTGCTACAGCGCTATCACCGTGGTGT 514
Qy      390 ThrGlnAenLeuArgProSerPheGlySerGlyAlaApsAlaAenIleAenIleValAlaGln 409
Db      513 ATGCAGACG-----AGCTGGGCC-----ATTGCTCT 487
Qy      410 PheSerSerArgGlyProThrIlyApsGlyAlaGlyIleApsProApsValMetAlaProGly 429
Db      486 TACTCTTAACACGCGCACCGTCTT-----GACATCTTGTGCTCCGGT 445
Qy      430 ThrTyrlleuSerAlaApsSerSerLeuAlaProApsSerSerPheTrpAlaAenHis 449
Db      444 ACCGGCGTCTCTCTGCC-----TGGTACACTAAC 415
Qy      450 ApsSerLeuGlyAlaTrpMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAps 469
Db      414 AACAGCGCCCAACAGCATTAACCGTACCTCCATGCGCCACTCCCACTGCTGCTT 355
Qy      470 ValAlaGln--LeuArgGlnHisPheValApsApsArgGlyIleThrProIlyProSer 488
Db      354 GTCCTTACCGGCATCTCTGTCAACGAGGTCTCC-----GGCGTACTGGCGTACCAAC 301
Qy      489 LeuLeuValAlaAlaLeuIleAlaGlyAlaAlaApsValGlyLeu--GlyTrpProAps 507
Db      300 TGGCTCAAGACCACTGCTCTGTGCAAGATCATCGGCAACTCCGCACTCCCCCAAC 241
Qy      508 -----GlyAasnGly 511
Db      240 CTGATCGGCACCAACGCG 223

RESULT 13
LOCUS   TA319G10P          532 bp    DNA       linear   GSS 13-DEC-2000
DEFINITION
T. brucei sheared genomic DNA clone 319g10, forward sequence,
ACCESSION
AL492464
VERSION  AL492464.1  GI:11867408
KEYWORDS
GSS.
SOURCE   Trypanosoma brucei
          Trypanosoma brucei
          Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
REFERENCE
1 (bases 1 to 532)
   Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Actin, R.,
   Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
   Melville, S.B., Rajandream, M.A. and Barrell, B.G.
   Direct Submission
   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
   project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
   Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
   nh@sanger.ac.uk
COMMENT
Constructed at the Institute for Genomic Research (TIGR),

```

```

FEATURES
    source
        1..532
        /organism="Trypanosoma brucei"
        /mol_type="genomic DNA"
        /strain="TRB927"
        /db_xref="taxon:5691"
        /clone="319g10"

ORIGIN
Alignment Scores:
Pred. No.:      2,54e-13      Length:      532
Score:          217.50      Matches:      60
Percent Similarity: 48.4%      Conservative: 29
Best Local Similarity: 32.6%      Mismatches: 68
Query Match:    6.7%      Indels:      27
DB:             11      Gaps:          6

US-10-784-870-4 (1-639) x TA319G10P (1-532)

Qy      298 ValPheGlnSerIleMetApsSerSerglyLeuGly----- 310
Db      2   GATATGCCAAGTAAATGTTGTGCGCAAGGGGGAAGATTCTTCAAGGGTGGCTGCC 61
Qy      311 -----GlyLeuProSerAenLeuGlnThrLeuPheSerGlnAlaPhe 324
Db      62 CATCCAGTCAAGAGCTTGTCTCCGCCAGACTTACTCAATATTATTCGCCCGGATAT 121
Qy      325 SerAlaGlyAlaArgIleHisThrAenSerTrpGlyAlaAlaValaasnGlyAlaTrpThr 344
Db      122 GGGCTGTAGAGCCGCTGTGTTCAACTCGTGGGGTTTGTGCTCCCTCCGATATTCT 181
Qy      345 ThrApsSerArgAenValApsAerGlyValAlaGlyA--ApsApsMetThrIleLeuPhe 363
Db      182 GCTGTGAAAGAGATGTGATGATGATGTTGCCAGTATGATGACATGCGTACTTATCTTC 241
Qy      364 AlaAlaGlyAasnGlyArgProAasnGlyThrIleSerAlaProGlyThrAlaApsAps 383
Db      242 TCCACTGGCAAGTATCCAAAGTGC-----CTATGATCTCGTGTGCTGTAAAGAC 295
Qy      384 AlaIleThrValGlyAlaThrGlnAenLeuArgProSerPheGlySerGlyAlaApsAps 403
Db      296 GTGATGTGGCTGGGGGTACACAAAGATGC-----TTTGACGCTTGAAAGAC--- 343
Qy      404 IleAenHisValAlaGlnPheSerSerArgGlyProThrIlyApsAerGlyAlaGlyLeuPro 423
Db      344 -----ATTGTTCTTCTGTTTCTTCTGCAATGTCGCAACATACGACGCTGAGTAAACCC 397
Qy      424 ApsValMetAlaProGlyThrTyrlleuSerAlaApsSerSerLeuAlaProApsSer 443
Db      398 GATCTGTGCGTCCCGGGAAGAGGTGTGTCTCTTCTCTGCGAAAGATCATGCT 457
Qy      444 SerPheTrpAlaAasnHisApsSerGlyAlaTrpMetGlyGlyThrSerMetAlaThr 463
Db      458 AAA-----CAATGTAAAGTGTGGCCCAAGCGAGTTCATCATGCGCAACT 502
Qy      464 ProIleValAla 467
Db      503 GCGGCGTGGCG 514

RESULT 14
LOCUS   CN808705          874 bp    mRNA       linear   EST 27-MAY-2004

```

```

Dy      419 GTCAATTCTGGTGTCTCTTCCTGCGCCGCCCTGGCAACAGTAACCCGGATGCCCAAGC 478
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      375 lIeSerAlaProGlYThrAlaLysAsnAlaIleThrValGIvalAthr-----GluAsn 392
          :::: :::: :::: ||||| ||||| :::: ::::
Db      479 ACCTCT---CCGCCTCCGAAGCCTTCTGCGACACTGTGTGTGCTCTTGCGGAAATGAC 535
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      393 LeuArgProSerPheGlySerTYzAlaAspAsnIleAsnIValAlaGlnPheSer 412
          :::: :::: :::: ||||| ||||| :::: ::::
Db      536 AGCCGATCTTCTCTTCCCACTAC----- 559
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      413 ArgGlyProThrlYlaAspGIyAlaGlyLeuVroAspValMetAlaProGlylThrTYlle 432
          :::: :::: :::: ||||| ||||| :::: ::::
Db      560 -----GGCAGAAT--GTCAATATTCTTGCTCTGTGAGCAATGTT 598
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      433 LeuSerAlaArgSerSerIleuAlaProAspSerSerPheTrpAlaAsnIlaAspSerLYs 452
          :::: :::: :::: ||||| ||||| :::: ::::
Db      599 CTTCACAC-----TCGATTGGT-----GGCCGC 622
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      453 TYzAlaTyMeClYglYThrSerMetAlaThrProIleValAlaGlyAsnValAlaGln 472
          :::: :::: :::: ||||| ||||| :::: ::::
Db      623 ACAACAACCATCTTGGAACCTCCACAGGATCTCCCAATATGCGCGGTCTGGCTGTAC 662
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      473 LeuArgGlnIlePheValIlyAsnArgTYlleThrProLYsProSerLeuLeuVala 492
          :::: :::: :::: ||||| ||||| :::: ::::
Db      683 CTCAGT-----GCGCTCCAAGCAAGACTACCCTCGCGCTTTCGAAAGAG 730
          :::: :::: :::: ||||| ||||| :::: ::::
Qy      493 AlaLeuIleAlaGlyAlaAlaAspValaGlyLeuGlyTYzProAsnGly 508
          :::: :::: :::: ||||| ||||| :::: ::::
Db      731 ATCCAGAACACTGTCTTCAACAAGACTGTCTCACCGGTGCTCTGTGAC 778
          :::: :::: :::: ||||| ||||| :::: ::::

```

DEFINITION	pnc 18302 Phytophthora infestans T30-4 Phytophthora infestans genomic, genomic survey sequence.
ACCESSION	CM803334
VERSION	CM803334.1
KEYWORDS	GSS.
SOURCE	Phytophthora infestans (potato late blight agent)
ORGANISM	Phytophthora infestans Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.
REFERENCE	1 (bases 1 to 555)
AUTHORS	Randall,T.A., Dwyer,R.A., Huitema,E., Beyer,K., Cvitanich,C., Kelkar,H., Fong,A.M., Gates,K., Roberts,S., Yatzkan,E., Gaffney,T. Law,M., Testa,A., Torto-Alalibo,T., Zheng,M., Zheng,L., Mueller,E. Windas,J., Binder,A., Birch,P., Gisi,U., Govers,F., Gow,N.A., Mauch,F., Van West,P., Waugh,M.E., Yu,J., Bolter,T., Kamoun,S., Lam,S.T. and Judelson,H.S.
TITLE	Large-scale gene discovery in the oomycete Phytophthora infestans reveals likely components of phytopathogenicity shared with true fungi
JOURNAL	Mol. Plant-Microbe Interact. 18 (3), 229-243 (2005)
PUBMED	15782637
COMMENT	Contact: Judelson HS Department of Plant Pathology University of California PLPA, Riverside, CA 92521, USA Tel: 9518274199 Fax: 9518274294 Email: howard.judelson@ucr.edu Class: shotgun.
FEATURES	Location/Qualifiers
SOURCE	1..555 /organism="Phytophthora infestans" /mol_type="genomic DNA" /isolate="T30-4" /db_xref="taxon:4787" /sex="A1" /tissue_type="total mycelium" /clone_id="Phytophthora infestans T30-4" /note="Vector: pUC18"

ORIGIN

Alignment Scores:

Pred. No.:	6,69e-13	Length:	555
Score:	214.00	Matches:	60
Percent Similarity:	50.7%	Conservative:	17
Best Local Similarity:	39.5%	Mismatches:	49
Query Match:	6.6%	Indels:	26
DB:	10	Gaps:	5

US-10-784-870-4 (1-639) x CW803334 (1-555)

```
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleThrAsnSerTyrGlyAlaAlaValAsn 340
    |||::|
Db 118 TCMAAGACTGCTCCAAAGCAGCAATATTGTTACACAGTTGGGT-----GGC 168

QY 341 GlyAlaTyrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThr 360
    |||::|
Db 169 GGTCAAGGTCTCAGATGTTGATGGGTTATCAAGCTTGGCAGGTGCGAGGATCATT 228
    |||::|
QY 361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
    |||::|
Db 229 CCGTGTAGCTCAGGGAATCTGCTCTTACTGCGTACTATTCATCGCTGCGAGAT 288

QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyr 400
    |||::|
Db 289 TCGGCTAGTGTGATCAAGTAGGTGCGACCGAT----- 321

QY 401 AlaAspAsnIleAsnHis--ValAlaGlnPheSerSerArgGlyProThrLysAspGly 419
    |||::|
Db 322 -----ATCAACGACGAGACTTGGCTCTTTCAGCAGCAAGGCCGACAGTCCAGGT 372

QY 420 ArgIleLeuProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeu 439
    |||::|
Db 373 CTCGAAACCGGATGGCAGCCCGGTGCTCTGGTATTG----- 414

QY 440 AlaProAspSerSerPheThrAlaAsnIleAspSerLysTyrAlaTyrMetGlyGlyThr 459
    |||::|
Db 415 -----TCATCTGCTGACGAGC--GACGCTCTTACTGCTTCAATCCGGCTCC 462

QY 460 SerMetAlaThrProIleValAlaGlyAsnValAla 471
    |||::|
Db 463 AGTATGGCTGCACGCGACGTGCGAGTGCATCGCT 498
```

Search completed: April 8, 2006, 07:51:36
Job time : 6643 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 03:16:46 ; Search time 520 Seconds
(without alignments)
2184.350 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267

Sequence: 1 MRKXKVLVLSAAALSTV.....EVOAYNVPGPNFSLATLV 639

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xip
-Q=/abs/ABSSWEB.epool/US10784870/runac.07042006_143420_27544/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cd1
-LIST=45 -DOCLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NCM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10784870.WCGN_1_1_512_rn1ac.07042006_143420_27544
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DB TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -YGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/68.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/68.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3267	100.0	1920	3	US-09-509-814A-3	Sequence 3, Appl1
2	3267	100.0	1920	3	US-09-920-954-3	Sequence 3, Appl1
3	3072	94.0	1923	3	US-09-509-814A-7	Sequence 7, Appl1
4	3072	94.0	1923	3	US-09-920-954-7	Sequence 7, Appl1
5	3066	93.8	1923	3	US-09-509-814A-5	Sequence 5, Appl1
6	3066	93.8	1923	3	US-09-920-954-5	Sequence 5, Appl1
7	2799	85.7	1977	3	US-08-873-479-41	Sequence 41, Appl1
8	477	14.6	1977	3	US-08-894-818B-2	Sequence 2, Appl1
9	477	14.6	1977	3	US-09-445-472-11	Sequence 11, Appl1

10	477	14.6	1977	3	US-10-090-624-11	Sequence 11, Appl1
11	477	14.6	1977	3	US-09-841-553-2	Sequence 2, Appl1
12	443	13.6	1962	3	US-08-894-818B-34	Sequence 34, Appl1
13	443	13.6	1962	3	US-09-445-472-15	Sequence 15, Appl1
14	443	13.6	1962	3	US-10-090-624-15	Sequence 15, Appl1
15	443	13.6	1962	3	US-09-841-553-34	Sequence 34, Appl1
16	443	13.6	1962	3	US-09-841-553-34	Sequence 34, Appl1
17	432	13.2	1977	3	US-09-902-540-959	Sequence 959, App
18	432	13.2	1977	3	US-08-894-818B-6	Sequence 6, Appl1
19	417.5	12.8	1977	3	US-09-841-553-6	Sequence 6, Appl1
20	417.5	12.8	1977	3	US-09-445-472-2	Sequence 2, Appl1
21	417.5	12.8	1977	3	US-10-090-624-2	Sequence 2, Appl1
22	417.5	12.8	1977	3	US-08-894-818B-4	Sequence 4, Appl1
23	362.5	11.1	2539	3	US-09-841-553-4	Sequence 4, Appl1
24	362.5	11.1	2539	3	US-09-514-340-3	Sequence 3, Appl1
25	357.5	10.9	2809	3	US-09-000-016-3	Sequence 1, Appl1
26	357.5	10.9	2809	3	US-09-514-340-1	Sequence 1, Appl1
27	326	10.0	2835	2	US-08-750-532-2	Sequence 2, Appl1
28	326	10.0	4765	2	US-08-750-532-8	Sequence 8, Appl1
29	326	10.0	4765	2	US-08-894-818B-7	Sequence 7, Appl1
30	326	10.0	4765	3	US-09-445-472-5	Sequence 5, Appl1
31	326	10.0	4765	3	US-10-090-624-5	Sequence 5, Appl1
32	326	10.0	4765	3	US-09-841-553-7	Sequence 7, Appl1
33	322	9.9	2532	2	US-07-671-376C-4	Sequence 4, Appl1
34	316.5	9.7	1306	3	US-09-966-921A-1	Sequence 1, Appl1
35	316.5	9.7	1330	3	US-09-966-921A-5	Sequence 5, Appl1
36	303.5	9.3	1170	3	US-09-902-540-8987	Sequence 8987, Ap
37	303.5	9.3	1859	3	US-08-894-818B-15	Sequence 15, Appl1
38	303.5	9.3	1859	3	US-09-841-553-15	Sequence 15, Appl1
39	292	8.9	898	2	US-08-750-532-7	Sequence 7, Appl1
40	283.5	8.7	564	2	US-08-750-532-11	Sequence 11, Appl1
41	283.5	8.7	564	2	US-08-894-818B-14	Sequence 14, Appl1
42	283.5	8.7	564	3	US-09-841-553-14	Sequence 14, Appl1
43	278.5	8.5	1191	2	US-08-434-553-3	Sequence 3, Appl1
44	278.5	8.5	1191	2	US-08-459-967-3	Sequence 3, Appl1
45	278.5	8.5	1191	2	US-08-460-327-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-509-814A-3
; Sequence 3, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..(1920)
; US-09-509-814A-3
Alignment Scores:

Prod. No.: 6,45E-311 Length: 1920
Score: 3267.00 Matches: 639
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-509-814A-3 (1-1920)

```
QY 1 MetArgLysIleValPheLeuSerValLeuSerIleAlaIleLeuSerThrVal 20
DB 1 ATGAGAAAGAAAGAGGTTTATCTGTTTATCAGCTGCAGCAATTCGTGCAGCTGT 60
QY 21 AlaLeuAsnAsnProSerIleArgIlePheAspLeuAspPheIle 40
DB 61 GCATTAAACATCCCTCGCTGGTGAAGCAAGGACTTGTGATCTGATTTAAAGAAATT 120
QY 41 GlnThrThrAspValSerGlyPheSerGlyIleArgIleThrGlyAlaIleAlaPhe 60
DB 121 CAACCAACACCGATGTCAGTGGTTTCTCCAAACAGCGAACAAAGGTGGCGCTGCATT 180
QY 61 LeuValGluSerGluAsnValIleLeuLeuIleGlyLeuLeuIleValIleVal 80
DB 181 CTGGTGAAGTCTGAAAATGGAACCTTTAAAGCAATTGCTAAAGAACTTGAAACAGTA 240
QY 81 ProAlaAsnAsnIleLeuIleIleValGlnPheAsnGlyProIleLeuGluIleThrIle 100
DB 241 CCGGCAAAATATAAATCCCATATGTCAAATTAAGGCCCACTTTAAGAAAGAAACAA 300
QY 101 GlnIleLeuGluIleThrThrGlyAlaIleValIleLeuAspThrIleProAspThrAlaIle 120
DB 301 CAGAACCTAGAGCAATCGAGCAAGATTCGTGACTACATCCCTGATTTAGCATATTT 360
QY 121 ValGluThrGluIleAspValGlnSerIleValIleArgSerIleGluIleValGluSerVal 140
DB 361 GTCGATATAGAGGGGATGTTCACTCAAAAGTCCGCTCCATGAAACAGTGGAAATCAGT 420
QY 141 GluProThrLeuProIleThrIleAspProGlnLeuPheThrIleGlyIleAspThr 160
DB 421 GAGCCATCTCTGCCAAATACAAATATGATCCCACTTTTCAAAAGGCCGATGAGAG 480
QY 161 LeuValIleAlaLeuAlaLeuAspThrIleGlnAsnAsnIleGluValGlnLeuArgIle 180
DB 481 CTGGTGAAGCGTTGGCGCTTGATACGAAGCAAGAACATTAAGAAATGCAATTAAGAGC 540
QY 181 IleGluIleIleAlaGlnIleThrValIleAspAspValIleIleThrIleThrAlaIlePro 200
DB 541 ATCGAAGAAATCGCTCAGTACGTAGCAAGCAATGACGTCATTATTTACGGCAAGCCCT 600
QY 201 GluThrIleValIleMetAsnAspValIleAlaArgGlyIleValIleValIleAspValIleGlnSer 220
DB 601 GAATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 221 SerThrGlyIleLeuThrGlyIleGlnIleValIleAlaValIleAspThrGlyLeuAspThr 240
DB 661 AGCTACGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 241 GlyArgAsnAspSerSerMetIleGluAlaPheArgGlyIleThrIleThrAlaIleValIle 260
DB 721 GGAAGAAAGCAAGTTCAGTACGTAGAGCTTCGGGATTAATTAACACACTATATGCA 780
QY 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyIleGlyThrIleIleValIleGlySer 280
DB 781 CTGGGTGCAACGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 281 ValLeuGlyIleAsnGlyIleAlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValIleGln 300
DB 841 GTATTATGAAATGCGCGCAACGAATTAAGAAATGCGCACTCAAGCGCAATCTGGTTTTCGA 900
QY 301 SerIleMetAspSerSerGlyIleLeuGlyIleLeuProSerAsnLeuGlnThrIlePhe 320
DB 901 TCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
```

```
QY 321 SerGlnAlaPheSerAlaGlyIleAlaArgIleIleIleThrAsnSerThrGlyIleAlaValIleAsn 340
DB 961 ACCCAAGCAATTCAGTGCAGTGCAGAAATTCATTAACAAATCTCTGGGGGGCGACGGTGAAT 1020
QY 341 GlyAlaIleThrThrAspSerArgAsnValIleAspAspThrValIleArgIleAsnAspMetThr 360
DB 1021 GGGGCTTCACACACAGATTCAGAAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 361 IleLeuPheAlaIleGlyAsnGluIleProAsnGlyIleThrIleSerAlaProGlyIleThr 380
DB 1081 ATTTCTTTTGGCGCTGGGAATGAAAGCCGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 381 AlaIleAsnAlaIleIleThrValGlyIleAlaThrGluIleAsnLeuArgProSerPheGlySerThr 400
DB 1141 GCTAAAGCAATTAACAGTGCAGTGCAGAAATTCGCGTCAAGCTCGGTTCTAT 1200
QY 401 AlaAspAsnIleAsnIleValIleGlnPheSerSerArgIleProThrIleAspGlyIleArg 420
DB 1201 GCAGATATATTAACACAGTTCACAGTTCCTTCCTGCGGCGCAACAAAGATGGCGCA 1260
QY 421 IleIleProAspValMetAlaProGlyIleThrIleLeuSerIleArgSerSerLeuAla 440
DB 1261 ATCAAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 441 ProAspSerSerPheThrAlaAsnIleAspSerIleThrAlaIleMetGlyIleThrSer 460
DB 1321 CCGATATCTCTCTCTGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 461 MetAlaThrProIleValIleAlaIleValIleAsnValIleGlnLeuArgGluIlePheValIleAsn 480
DB 1381 ATGGCAACACCGATGTTGGCGGGAATGTTGCAAGCTCCGAGACATTTTGTGAAAAT 1440
QY 481 ArgGlyIleThrProIleProIleProIleLeuIleValIleAlaIleGlyIleAlaIleAsp 500
DB 1441 AGAGCAATCACTCTTAAGCTCTCTTAATGAAGCAAGCTTGAATGTCAGAGTCTCTAT 1500
QY 501 ValGlyLeuGlyIleThrProAsnGlyIleAsnGlyIleThrIleValIleThrLeuAspIleSer 520
DB 1501 GTTGAATGGTATTCGCAACGAAACCAAGATGGGCGGAGTACCTGATTAATTCG 1560
QY 521 LeuAsnValIleAlaIleValIleAsnGlyIleSerIleLeuSerIleThrIleValIleThrIle 540
DB 1561 TTGAAGTGTGCTTATGTCAGAAATCCAGTGCCTATCAATGCAAAAGCGCATAT 1620
QY 541 ThrPheThrAlaIleAlaGlyIleProIleLeuIleIleSerLeuValIleThrAspAlaPro 560
DB 1621 ACCTTTACTGCAACGGCGGCAAGCATTTGAATACTCTCGTATGATGATGATGATGATGATGAT 1680
QY 561 AlaSerThrThrAlaSerValIleThrLeuValIleAsnAspLeuAspLeuValIleThrAlaPro 580
DB 1681 GCAGCACTACTGCTTCTGTAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 581 AsnGlyThrArgIleValIleGlyIleAsnAspPheSerAlaProPheAspAsnAsnThrAspGly 600
DB 1741 AACGCAACAAATATGTCGGGATGATCTTCAGACCAATTTGCAATTAACGGAATGGC 1800
QY 601 ArgAsnAsnValIleGluAsnValIlePheIleAsnSerProGlnSerGlyIleThrIleGlu 620
DB 1801 CGCAATTAACGTAAGAAATGATTTATTAATTCGCCCCCAAGTGAACATATATCACTTGAAG 1860
QY 621 ValGlnAlaIleThrAsnValIleProValGlyIleProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1861 GTGCAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1917
```

RESULT 2

US-09-920-954-3

; Sequence 3, Application US/09920954

; Patent No. 6759228

; GENERAL INFORMATION:

; APPLICANT: TAKAIWA, MIKIO

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHIISA

; APPLICANT: KUBOTA, HIROMI

```

1  APPLICANT: HITOMI, JUN
2  APPLICANT: KAGEYAMA, YASUSHI
3  APPLICANT: SHIKATA, SHITSUM
4  APPLICANT: NOMURA, MASAFUMI
5  TITLE OF INVENTION: ALKALINE PROTEASE
6  FILE REFERENCE: 0327-0832-0PCT
7  CURRENT APPLICATION NUMBER: US/09/920,954
8  CURRENT FILING DATE: 2001-08-03
9  PRIOR APPLICATION NUMBER: 09/509,814
10 PRIOR FILING DATE: 2000-04-06
11 PRIOR APPLICATION NUMBER: PCT/JP98/04528
12 PRIOR FILING DATE: 1998-10-07
13 PRIOR APPLICATION NUMBER: JP 9-274570
14 PRIOR FILING DATE: 1997-06-08
15 NUMBER OF SEQ ID NOS: 24
16 SOFTWARE: PatentIn version 3.0
17 SEQ ID NO 3
18
19 LENGTH: 1920
20
21 TYPE: DNA
22
23 ORGANISM: Bacillus sp.
24
25 FEATURES:
26
27 NAME/KEY: CDS
28
29 LOCATION: (1)..(1920)
30
31 OS-09-920-954-3

```

QY 561 AlasSerThrThraIasSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
DB 1681 GCAGPACTCTCTCTCTCTGTAACCTGGTCAAGATGGATTTGGATTCATTACAGCACCA 1740
QY 581 AsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTTPAspGly 600
DB 1741 AACGGAACAAGATATGTCGGGAATGACTTCTCAGCACCATTTGACATATACTGGATGGC 1800
QY 601 ArgAsnAsnValGluLeuValPheIleAsnSerProGlnSerGlyThrTyrIleGlu 620
DB 1801 CGGAATAACCTTAAGATAATGATTAATTCGCCCAAGTGAACATATACCATTTGAG 1860
QY 621 ValGluAlaIleTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1861 GTCCAAAGCATATAATGTCGGTTGGACACAAATCTTCGTGGCATTTGTGAAC 1917
RESULT 3
US-09-509-814A-7
Sequence 7, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-509-814A-7
Alignment Scores:
Pred. No.: 9, 38e-292
Score: 3072.00
Percent Similarity: 97.0%
Best Local Similarity: 93.4%
Query Match: 94.0%
DB: 3
Gaps: 0
US-10-784-870-4 (1-639) x US-09-509-814A-7 (1-1923)
QY 2 ArgGlyValLeuValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
DB 7 AAGAGAAAGAGAGTGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGCACTGTGGC 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaIleThrPheAspLeuAspPheLeuGlyIleGln 41
DB 67 TTAAAGTAATCATCTGACAGTGTGTCAGAGAAATTTTGATCTGGATTTCAAGAAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerLeuGlnArgGlnThrGlyValAlaIlePheLeu 61
DB 127 ACAACAACCTGATGCTTAAGGTTTCTTCAGACAGGGCAGACTGGTGCTCTGCTTTTCTG 186
QY 62 ValGluSerGluAsnValIleLeuLeuLeuGlyLeuLeuValLeuLeuGluThrValPro 81

DB 187 GTGGAATCTGAATAATGTGAACCTCCCAAAAGGTTTGACAGAGAGCTTGAAACATGCCG 246
QY 82 AlaAsnAsnValLeuValIleValAlaGlnPheAsnGlyProIleLeuGluGluThrIleGln 101
DB 247 GCAAATTAATAAATCTCATATTAATCCAAATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGluThrThrGlyValAlaValIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
DB 307 CAGCTGGAAAAACAGGGGCAAAAGTTCTCGACTACATCTGATTAATGCTTACATTGTC 366
QY 122 GlnTyrGluGlyAspValGlnSerLysValArgSerIleGluIleValGluSerValGlu 141
DB 367 GAGTATGAGGGCGAGTTAAGTCAGAACACACCATTTGACAGCTGGAAATCCGGAGG 426
QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyValAsnThrLeu 161
DB 427 CCTTATTTGCCATATACAGAAATGATCCCGCTTTTCACAAAGGGGCAACAGAGCTT 486
QY 162 ValLysAlaLeuValLeuAspThrLysGlnAsnAsnValGluValGlnLeuArgGlyIle 181
DB 487 GTAAAGACGTGGCGCTTGATACAAAGCAAAATTAAGAGTGCATTAAGAGCATC 546
QY 182 GlnGluIleAlaGlnTyrValAlaSerAsnAspValIleTyrIleThrAlaLysProGlu 201
DB 547 GAACAATCGCAAAATTCGCAATTAAGCAATGATGCTATATATTACGCCAAAGCCTGAG 606
QY 202 TyrIleValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
DB 607 TATTAAGGTGATAAAGATGTTGCCGGTAATTCGCAAGCGAGTGGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
DB 667 TACGGGTGTATGACAAAGACAGATCGTAGCGGTGCCGATACAGGGCTTGATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
DB 721 CGCAATGACAGTTCATGATGATGATGACCTTCGGCGGAAATTAATCGCATTAATGATGAT 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
DB 787 GCACGAGAAATGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
DB 847 TTAGAAACGGCTCCATTAATTAAGAAATGGCGCTTCAGCGAAATCTTGCTTCCAACTT 906
QY 302 IleMetAspSerSerGlyLysGlyLysGlyLysLeuProSerAsnLeuGlnThrLeuPheSer 321
DB 907 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTTPGlyAlaAlaValAsnGly 341
DB 967 CAAGCATTCAGTGTGGGCGCAAGTTCATCAAACTCTGGGGGAGCAGTGAATGGG 1026
QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
DB 1027 GCTTACAAACAGATTCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyIleThrIleSerAlaProGlyThrAla 381
DB 1087 CTTTTCGTGGCGGAAATGAAGAACCGAAGCGGGAACCATCAAGTGCACAGGCACT 1146
QY 382 LysAsnAlaIleThrValGlyAlaThrGluLeuLeuAspProSerPheGlySerTyrAla 401
DB 1147 AAAAATGAATACAGTGGAGCTACGGAACCTCCCAAGCTTTGGGCTTATGCG 1206
QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyValArgIle 421
DB 1207 GACAAATATCAACCAATGTCACAGTTCCTTCACTGGAACGCAAAAGATGAGAGGATC 1266
QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
DB 1267 AAACCGAGTGTGATGACCGGGAAAGCTTCATATCATGACAAATCTTCTCTTGACACG 1326


```
QY 442 AppSerPheThrAlaAsnHisAppSerIleuValAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTTCTGGGGAAACATGACAGTAAATATCATATGAGTGAACCTCATG 1386
QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIlyAsnArg 481
Db 1387 GCTAACCCGATCGTGTGTGAACGTGGCAAGCTTGATGAGCATTTTGTGAAGAAAACAGA 1446
QY 482 GlyIleThrProIlyPProSerIleuLeuValAlaLeuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCATACACCAAGCCCTTCTATTAAGCCGACATGATGCGGTGACGCTGACATC 1506
QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspIlySerLeu 521
Db 1507 GGCCTTGCTACCCGAAACGATACCAAGATGGGAGAGATGATGATTAATATCCCTG 1566
QY 522 AsnValAlaTyrValAsnGlySerSerAlaLeuSerThrSerGlnValAlaThrTyrThr 541
Db 1567 AACGTTGCCATATGAAAGAGTCCAGTTCTCTATCCACAGCCAAAAGGACGTAATCG 1626
QY 542 PheThrAlaThrAlaGlyIlyPProLeuIlyIleSerLeuValTyrSerAspAlaProAla 561
Db 1627 TTTACTGCTACTCGCCGCAAGCCTTTGAAAATCTCCCTGATGCTGATGCGCCCTGCG 1686
QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCAACAAGCTCTCCGACGCTGTGCAATGATCTGACCTGTGCATTTACCGCTCCAAAT 1746
QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTPAspGlyArg 601
Db 1747 GGCACACAGTATATAGAAATGACTTACTTCCCATACATATGATATGAGATGCGCGC 1806
QY 602 AsnAsnValGlnAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGlnVal 621
Db 1807 AATTAAGTGAAGAAATGATATTATTAATGACACCAAGCGGACGTATCAATTGAAGTA 1866
QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerIleuAlaIleValAsn 639
Db 1867 CAGGATTAATACGATACCGGTTGACACACAGAACTTCTCGTTGCAATTGAAAT 1920

RESULT 4
US-09-920-954-7
; Sequence 7, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
```

```
US-09-920-954-7
Alignment Scores:
Pred. No.: 9,38e-292 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-7 (1-1923)
QY 2 ArgIlyLeuValPheLeuSerValIleuSerAlaAlaIleuSerThrValAla 21
Db 7 AAGAGGAAAGAGGATTTTATCTGTTTATTCAGCTGACGCAATTTGTGACTGTGCG 66
QY 22 LeuAsnAspProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIlyGlnIleGln 41
Db 67 TTAAGTATTCATCTCAGGTGTCAGAGGAATTTTGATCTGATTTCAAGGAATTCAG 126
QY 42 ThrThrThrAspValSerGlyPheSerIlyGlnArgGlnThrGlyAlaAlaIleLeu 61
Db 127 ACAACAACGATGCTAAGGTTTCTCCAAAGCAGGCGCAGACTGTGCTGCTTTTCTG 186
QY 62 ValGluSerGlnAsnValIlyIleLeuLeuGlyGlyLeuLeuValIleGlnIlyThrValPro 81
Db 187 GTGGAATCTGAATATGTAATGAACTCCCAAGGTTTGCAAGAGAGCTTGAAACGATCCG 246
QY 82 AlaAsnAsnIlySerLeuHisIleValGlnPheAsnGlyProIleLeuGlnIlyThrIlyGln 101
Db 247 GCAAATATATTAATCTCATATTAATCAATTAAGCAATTTTAAGAGAAACAAACAG 306
QY 102 IlyLeuGlnIlyThrGlyAlaIlyIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGGAAGAAACAGGCGCAAGATTTCTGCACTACATACCTGATATGCTTCAATGTC 366
QY 122 GlnTyrGlnGlyIlyAspValGlnSerIlyValArgSerIleGlnHisValGlnSerValGln 141
Db 367 GAGTATGAGGCGCATTTAATGACCAACAGCAACCAATTAAGCACTGGAATCCGTGAG 426
QY 142 ProTyrIleuProIlyIlyIleAspProGlnIleuPheThrIlyGlyAlaSerThrLeu 161
Db 427 CCTATTTCGCCATATATACAGAAATGATCCCGACTTTTCAAAAGGCGATCAGAGCTT 486
QY 162 ValIlyAlaLeuAlaLeuAspThrIlyGlnAsnAsnIlyGlnValGlnLeuArgGlyIle 181
Db 487 GTAAAGACGATGCGCTGATACAAAGCAGAAATAAAGAGGTGCAATTAAAGGACATC 546
QY 182 GlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaIlyProGln 201
Db 547 GAACAATGCAATTCGCAATTAACATGATGATGCTATATTAATTCGGCAAGCCCTGAG 606
QY 202 TyrIlyValIleAsnAspValAlaArgGlyIleValIlyAlaAspValAlaGlnSerSer 221
Db 607 TATAGGTATGATATATATTTGCGGTGCAATTTGTCAAGGCGATGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnIlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTATGCAACAGCAAGATCGTAGCGTTGCCGATCAAGGCTTGATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGlnAlaPheAspGlyIlyIleThrAlaLeuTyrAlaLeu 261
Db 727 CGCAATGACAGTTCATGATGATGAGCTTCCCGGGGAAATTAATGCAATTATATGCAATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGAGCAATTAATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 282 LeuGlyAsnGlyAlaThrAsnIlyGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
Db 847 TTAGAAACGCGCTCCATTAATTAAGGAATGGCGCTCAGGCGCAATCTAGCTTCAATCT 906
QY 302 IleMetAspSerSerGlyIlyLeuGlyIlyLeuProSerAsnLeuGlnThrLeuPheSer 321
```



```
Db 547 GAAACAATCCGACAAATTCGCAATGACATGATGCTATATATATGAGGCAAGCCCTGAG 606
Qy 202 TTTTCTGCTACTGCGCGCAAGCCTTTGAAATCTCCCTGATGCTGATGACCCCTGCG 1686
Db 607 TATTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1746
Qy 222 TTTTCTGCTACTGCGCGCAAGCCTTTGAAATCTCCCTGATGCTGATGACCCCTGCG 1806
Db 667 TACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1806
Qy 242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyLeuValIleThrAlaLeuValIle 621
Db 727 CGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1866
Qy 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyValIleGlyThrHisValAlaGlySerVal 639
Db 787 GGAACGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920
Qy 282 LeuGlyAsnGlyValAlaThrAsnValGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
Db 847 TTAGGAAACGCTCCGACCTAATTAAGGAATGCGCTCAGGCAATCTGATCTTCCAAATCT 906
Qy 302 IleMetAspSerSerGlyLeuGlyValLeuProSerAsnLeuGlnThrLeuPheSer 321
Db 907 ATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Qy 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGly 341
Db 967 CAAAGATACAGTCTGCTGCGCAAGATTCATACAACTCTGCGGAGCAGCAGATGATGAGG 1026
Qy 342 AlaIleThrThrAspSerAlaGlnValAspAspTrpValArgValAsnAspMetThrIle 361
Db 1027 GCTTACCAACAAGATTCGAAATGATGATGATGATGATGATGATGATGATGATGATG 1086
Qy 362 LeuPheAlaAlaGlyAsnGlyValArgProAsnGlyValThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCCTGCTCCGGGAATGAGCAACGAGCGGAAACCATCACTGACCAAGCAGCAGCT 1146
Qy 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTrpAla 401
Db 1147 AAAAATGCAATTAACAGTCGAGCTACGGAACCTCCGCGCAAGCTTTGGGTCTTAAGCG 1206
Qy 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyValIle 421
Db 1207 GACAAATATCAACATGTCGACAGTTCTCTTCACTGTCAGCAAGGATGAGCGGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTrpIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGGATGTCATGCGAGCGGAAACGTTCACTATCAACAGATCTTCTTGGACCG 1326
Qy 442 AspSerSerPheThrAlaAsnHisAspSerLeuTrpAlaTrpMetGlyValThrSerMet 461
Db 1327 GATTCCTCTCTTGGGGAACATGACATGATTAATGATCAATGAGGTGGAACGTCATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValLysAsnArg 481
Db 1387 GCTAACCGGATGTCGTCGAGAAACGTGCGACAGCTTGAGCACTTTTGGAAAAACAGA 1446
Qy 482 GlyIleThrProLysProSerLeuLeuValAlaIleLeuIleAlaGlyAlaAlaAspVal 501
Db 1447 GGCATTCACACCAAGCTTCTCTATTAAGCGGACATGATGCGCGGTGACGTCAGATC 1506
Qy 502 GlyLeuGlyValProAsnGlyAsnGlnGlyTrpGlyValThrLeuAspLysSerLeu 521
Db 1507 GGCCTTGGCTACCGGAACGTTAACCAAGATGGGGAACGATGACATTTGATTAATCCCTG 1566
Qy 522 AsnValAlaIleValAlaGlnLysSerSerAlaLeuSerThrSerGlnLysAlaThrTrpThr 541
Db 1567 AAGGTGCTTAATGAGCAAGTCCAGTTCTCTATCCACAGCAAGCAAGCAAGCAAGTACG 1626
Qy 542 PheThrAlaThrAlaGlyLysProLeuValIleSerLeuValTrpSerAspAlaProAla 561
```

```
Db 1627 TTTACTGCTACTGCGCGCAAGCCTTTGAAATCTCCCTGATGCTGATGACCCCTGCG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspValIleThrAlaProAsn 581
Db 1687 AGCACAACGCTCTCCGTAACGCTTGTCAAGATCTGACCTTGTCAATTAACGCTCCAAAT 1746
Qy 582 GlyThrArgTrpValGlyAsnAspPheSerAlaProPheAspAsnMetTrpAspGlyArg 601
Db 1747 GGCACACAGTATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1806
Qy 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTrpIleGluVal 621
Db 1807 AATTAAGTAAAGAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1866
Qy 622 GlnAlaIleAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1920

RESULT 6
US-09-920-954-5
; Sequence 5, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MUKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUHIKA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-5

Alignment Scores:
Pred. No.: 3,656-291 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-5 (1-1923)
Qy 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAAAGAAAGAGTGTATTTATCTGTTTATGCTGACGACATTTTGTGCTGCTGCG 66
Qy 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIleGln 41
Db 67 TTAAGTAATCACTCAGTGTGTCAGAGAAATTTGATCTGATTTCAAGGAATTCAG 126
Qy 42 ThrThrThrAspValSerGlyPheSerLysGlnArgGlyThrGlyAlaAlaPheLeu 61
Db 127 ACAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 186
```

QY 62 ValGluSerGluAsnValLysLeuLeuGlyLeuLeuLysLeuGluThrValPro 81
 DB 187 GTGGATCTGAAATATGAAATCTCCAAAGGTTTCAGAGAGAGCTTGAACAGCTCCG 246
 QY 82 AlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGluGluThrLysGln 101
 DB 247 GCAAATATTAATCACTCATATTCATCAATGAGCAATTTTAAAGAAACAAACAG 306
 QY 102 LysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
 DB 307 CAGCTGGAAAAACAGGGCAAGATTCAGATACCTGATATCATTCATTCATTCATTC 366
 QY 122 GluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGluSerValGlu 141
 DB 367 GAGTATGAGGCGAATTTTAAAGTACAGCAACAGCAATGAGCACTGGAAATCCGAGAG 426
 QY 142 ProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyLysSerThrLeu 161
 DB 427 CCTTATTTGCCGATATACAGAAATAGATCCCGAGCTTTTCAAAAGGGGCACTCAGAGCTT 486
 QY 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGlyIle 181
 DB 487 GTAAAAAGCGTGGCGCTGATACAAAGCAAAAAATAAAGAGGTGCAATTAAGAGGCATC 546
 QY 182 GluGluIleAlaGlnTyrValAlaSerAsnAspValIleTyrIleThrAlaLysProGlu 201
 DB 547 GAAACAATGCAACATTCGCATTAAGCAATGATGCTCTATATTTACGGCAAAAGCTTGAG 606
 QY 202 TyrLysValIleAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
 DB 607 TATAAGGTGATGATGATGTTGGCGGTGAATTTGCAAAAGCGGAGTGGCTCAGAGCAC 666
 QY 222 TyrGlyLeuTyrGlyGlnGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
 DB 667 TACGGGTGTATGACAAAGCAAGATCGTACGGTTCGATACAGGGCTTGAATCAAGCT 726
 QY 242 ArgAsnAspSerSerMetHisGluAlaPheArgGlyValIleThrAlaLeuTyrAlaLeu 261
 DB 727 CGCAATGACAGTTCATGATGATGAGCTTCCGGGAAATATTCGATTAATATGATTCG 786
 QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerVal 281
 DB 787 GAGCGGACGAATATAGCCAAATGATAGATGTCATGATGATGATGATGATGATGATG 846
 QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
 DB 847 TTAGGAAACGGCTCCACTATATTAAGGAATGGCGCTCAGCGCAATCTACTCTTCCATCT 906
 QY 302 IleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
 DB 907 ATCATGATATAGCGGTGGGAGACTTGGAGGACTTCTTCAAACTGTCAAACTTATTCACG 966
 QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisIleThrAsnSerTyrGlyAlaAlaValAsnGly 341
 DB 967 CAAGCATACAGTGGTGGTCCAGAAATTCATCAAACTCTGGGAGGACACATGATGGG 1026
 QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
 DB 1027 GCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATG 1086
 QY 362 LeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
 DB 1087 CTTTTCGCTGGCGGAAATAGAGACCGAAAGCGGAAACATCATGAGCAACAGGCAAGCT 1146
 QY 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
 DB 1147 AAAAATGCAATACAGTGGAGCTACGAAAACTCCGCGCAAGCTTGGCTTATGGG 1206
 QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421
 DB 1207 GACAAATATCAACCAATGTGGCAGATGTTCTTCACTGAGCACCAAAAGATGAGAGGATC 1266

QY 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
 DB 1267 AAACGGATGTATATGGCACCGGAAAGCTTCAATCATATCAGCAAGATCTTCTTGCACCG 1326
 QY 442 AspSerSerPheTyrAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
 DB 1327 GATTCCTCTCTTGGGGCGAACCATGACAGTAATATGATACATGATGGTGGAAAGCTCATG 1386
 QY 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValLysAsnArg 481
 DB 1387 GCTACACCATCTGTTGCTGGAAACGTGGCACGCTTCGTGAGCATTTTGTGAAAAACGA 1446
 QY 482 GlyIleThrProLysProSerLeuLeuLysAlaIleValIleAlaGlyAlaAlaAspVal 501
 DB 1447 GGCATCACCAACAGCTTCTCTATTAAGGGGCACTGATTTGGCGGTGAGCTGACATC 1506
 QY 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu 521
 DB 1507 GGCCTTGGCTACCGAACGGTAAACAGGATGGGAGCGAGTGAATGGATTAATCCCTG 1566
 QY 522 AsnValAlaTyrValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
 DB 1567 AACGTTGCTATGTAGAACAGATCCAGTCTCTATTCACACAGCCAAAAAGCGACGTACCG 1626
 QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrPheAspAlaProAla 561
 DB 1627 TTTACTGCTACCTGGCGAACGCTTTGAAATCTCCCTGGTATGATGATGATGATGATG 1686
 QY 562 SerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
 DB 1687 AGCACAACTGCTTCGTAACGCTTGTCAATGATGTGACCTTGTCAATACGCTCAAT 1746
 QY 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArg 601
 DB 1747 GGCACACGATGTAGAAATGACTTACTTCTCGCATACAAATGATGATGATGATGATG 1806
 QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
 DB 1807 AATACGTAGAAATGATTTTATTTATGACACCAACAGGAGCATATACATTAAGGTA 1866
 QY 622 GlnAlaTyrAsnValProValGlyProGlnAsnAspSerLeuAlaIleValAsn 639
 DB 1867 CAGGCTTAATACGTACCGGTTGAGCACACAGACCTTCTGTTGGCAATGTGAAT 1920

RESULT 7
 US-08-873-479-41
 ; Sequence 41, Application US/08873479
 ; Patent No. 5891701
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; TITLE OF INVENTION: Nucleic Acid Encoding A Polypeptide
 ; TITLE OF INVENTION: Having Protease Activity
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 5891701 No. 5891701disk of No. 5891701th America
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/873,479
 ; FILING DATE: 12-JUN-1997
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Agilis, Cheryl H
 ; REGISTRATION NUMBER: 34,086

REFERENCE/DOCKET NUMBER: 5251.000-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 TELEX:
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3003 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 US-08-873-479-41

Alignment Scores:
 Pred. No.: 1.27e-264 Length: 3003
 Score: 2799.00 Matches: 541
 Percent Similarity: 92.4% Conservative: 52
 Best Local Similarity: 84.3% Mismatches: 45
 Query Match: 85.7% Indels: 4
 DB: 2 Gaps: 2

US-10-784-870-4 (1-639) x US-08-873-479-41 (1-3003)

Qy 1 MetArglyS-----LysLysValPheLeuSerValLeuSerAlaAlaIleLeu 17
 Db 846 ATGAGAAAGAAAGATCGAAGAGGGTTTATCCGTTTATCATGCTGCACATATG 905
 Qy 18 SerThrValAlaLeuAenAenProSerAlaGlyAspAlaArgThrPheAspLeuAspPhe 37
 Db 906 TCTTCGTCCTTAAACAGCTCTTACTATGCGGCAACATTTGAAATGCACTTT 965
 Qy 38 LysGlyYIleGlnThrThrThrAspValSerGlyPheSerIysGlnArgGlnThrGlyAla 57
 Db 966 AAGGGAGTAAACACTTACGCTAGAGAGGCTGCCCAAGCAAGCAAGAAACCGGGAAG 1025
 Qy 58 AlaAlaPheLeuValGluSerGluAenValLysLeuLysGlyLeuLeuLysValLeu 77
 Db 1026 GCATCTTTCTTGTAACTCTGAATAATGTAAATCCCAAGACTATTCAAAAGAACTA 1085
 Qy 78 GluThrValProAlaAenAenLysLeuHleIleValGlnPheAenGlyProIleLeuGlu 97
 Db 1086 GAAGTAGTTCAGCGGATTAACAAGCTATATATGCTTCAATTGACGCACTATTGTAAG 1145
 Qy 98 GluThrLysGlnLysLeuGlnThrThrGlyAlaLysIleLeuAspYrIleProAspYr 117
 Db 1146 GAAACGCACTTCAACTAGAGAGACCGGAGCGGAAATCTCGATTACATACCGAGATTAC 1205
 Qy 118 AlaThrIleValGluGlyArgGluAspValGlnSerLysValArgSerIleGluHleVal 137
 Db 1206 GCTTATATGTGCAATATGATGGGAGTAAAGCGCTAACTAAAGCAATTCGCAATTG 1265
 Qy 138 GluSerValGluProLysLeuProLysYrIleLysIleAspProGlnLeuPheThrLysGly 157
 Db 1266 GAATCGGTGACCATATTACCTTATATTAATTAACCCGCAATTATTTCCAGAGCA 1325
 Qy 158 AlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAenAenLysGluValGln 177
 Db 1326 GCTTCGAATTAAGTAGAACAAGTAGCTTGAATTAAGCAAGAAAGTAAGAGTAGCT 1385
 Qy 178 LeuArgGlyIleGluGluIleAlaGlnYrValAlaSerAenAspValHleSerYrIleThr 197
 Db 1386 TTAAAGGATTTGAACAATTTGCCAATATGCGCAACAATATGATGATTAATACGTAAAC 1445
 Qy 198 AlaLysProGluLysValLysValLysValLysValLysValLysValLysValLysVal 217
 Db 1446 CCAAGCGCTGAATTCGAAGTTTGAATGAGTGGCCGCTGCACTTGTGAAGCAAGACGTC 1505
 Qy 218 AlaGlnSerSerYrGlyLeuYrGlyGlnGlnIleValAlaValAlaAspThrGly 237
 Db 1506 GCAACAATAATCACTTGCTTATATGACAAGGACAGATTGTACAGATTGCTGATATCGGG 1565
 Qy 238 LeuAspThrGlyArgAenAspSerSerMetHleGluAlaPheArgGlyLysIleThrAla 257

Db 1566 CTGATATCAGAGAAATATGACAGTGCATGCAATGAAACATTCGCCGCTAATATACCGCA 1625
 Qy 258 LeuYrAlaLeuGlyArgThrAenAenAlaAenAspThrAenGlyHleGlyThrHleVal 277
 Db 1626 CTATATGCACTGGGCGAGAAAGAAATTAACGCAATGATCCAATGGAATGGAACCAATGTT 1685
 Qy 278 AlaGlySerValLeuGlyAenGlyAlaThrAenLysGlyMetAlaProGlnAlaAenLeu 297
 Db 1686 GCTGGATCTGTGTAGGAAT---GCTACAAATTAAGGATGGACCGCAAGCAACCACTA 1742
 Qy 298 ValPheGlnSerIleMetAspSerSerGlyYrLeuGlyYrLeuProSerAenLeuGln 317
 Db 1743 GCTTTCAATCTATATGATAGTATGCTGGAGGCGCTGGAGAGACTACTCTGATATCTACA 1802
 Qy 318 ThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHleThrAenSerTrpGlyAla 337
 Db 1803 ACATATTCAGTCAGCAATATAGTGTGAGCGCAAGAAATTCATACGAATTCATGGGGGGCT 1862
 Qy 338 AlaValAenGlyValYrThrThrAspSerArgAenValAspAspYrValArgLysAen 357
 Db 1863 CCAGTAAACGCTGCTATAGACAGACTCTCGAAATGTTGATATTAATGAGAAAAAT 1922
 Qy 358 AspMetThrIleLeuPheAlaIleGlyAenGlyLysArgProAenGlyYrThrIleSerAla 377
 Db 1923 GATATGACATTTCTTTTGGCGCGCAATGAGGACCAAGTACCGGTATCATTCAGTGCA 1982
 Qy 378 ProGlyThrAlaLysAenAlaIleThrValGlyAlaThrGluAenLeuArgProSerPhe 397
 Db 1983 CCAAGAAACAGCAAAAAATGCGATTACAGTTGCGGCAACGAAACCTAACGTCACAGCTTC 2042
 Qy 398 GlySerYrAlaAspAenIleAenHleValAlaGlnPheSerSerArgGlyProThrLys 417
 Db 2043 GGATCTTAATCGGATATATTAACAATGTTCAATTCCTTCAACAGAGCTCTACTACGA 2102
 Qy 418 AspGlyYrGlyLysProAspValMetAlaProGlyYrThrYrIleLeuSerAlaArgSer 437
 Db 2103 GATGAGATTTAAAGCGAGCTCATGACCAAGGATCGTATATTCCTGCTGACAGTCA 2162
 Qy 438 SerLeuAlaProAspSerSerPheTrpAlaAenHleAspSerLysYrAlaYrMetGly 457
 Db 2163 TCATTTAGCTCAGATTCCTCATTCCTGGGCAACCAAGATGTAATATGCTCATATGCGT 2222
 Qy 458 GlyThrSerMetAlaThrProIleValAlaGlyAenValAlaGlnLeuArgGluHlePhe 477
 Db 2223 GGTACTTCTATGCTACTCTCAATGTGACGAGTAACTTGCCACAATTAAGGGAGCACTTT 2282
 Qy 478 ValLysAenAspGlyIleThrProLysProSerLeuLysValAlaLeuIleAlaGly 497
 Db 2283 GTAAATAATTAAGGGGTAACCTCTTAAGCTTCCCTTTTAAACCTGCTTAAATTCAGGT 2342
 Qy 498 AlaAlaAspValGlyLeuGlyYrProAenGlyAenGlnGlyYrProGlyArgValThrLeu 517
 Db 2343 GCTGGGAGTGTGCACTTGCTTCCAAATGTAATGTAACCAAGAGGGGAAGATGACTTA 2402
 Qy 518 AspLysSerLeuAenValAlaYrValAenGluSerSerAlaLeuSerThrSerGlnLys 537
 Db 2403 GATAAATCCCTAAATGCTGATTTGGAATGAACAGAGCCCTTATCAACAACTCAAAA 2462
 Qy 538 AlaThrYrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrpSer 557
 Db 2463 GCAACATATTCGTTACGCTCAAGCTGGTAACCTTTAAATAATATCACTGTTGGTCA 2522
 Qy 558 AspAlaProAlaSerThrThrAlaSerValThrLeuValAenAspLeuAspLeuValIle 577
 Db 2523 GATGACCAAGTAGACAGAGGATCACTTAATGTAATGATTAATGATTAATGATTAATG 2582
 Qy 578 ThrAlaProAenGlyYrArgYrValGlyAenAspPheSerAlaProPheAspAenAen 597
 Db 2583 ACTGACCAAAATGGAATCAATATCGTCGGAATAATGACTTTACAGCAACGATATGATTAACAT 2642
 Qy 598 TrpAspGlyArgAenAenValGluAenValPheIleAenSerProGlnSerGlyThrYr 617

Db 2643 TGGGATGGAGAAACACGTGAGAAAATGTTTATCATGCTCTCAAGCGGAGTAT 2702
Qy 618 ThrileglValAlGlnAlaIleTyrAenValProValGlyProGlnAsnPheserleuAlaIle 637
Db 2703 ACGGTGAGGTGACGCTTACATGATACAGTACGAGTCCGCAACCTTTCTTACGAGATT 2762
Qy 638 ValAsn 639
Db 2763 GTACAT 2768

RESULT 8
US-08-894-818B-2
Sequence 2, Application US/08894818B
Patent No. 6261822
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSER: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,818B
FILING DATE: 20-MAY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03253
FILING DATE: 07-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TAKAKURA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1977 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-894-818B-2

Alignment Scores:
Pred. No.: 9.76e-37 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.84 Conservative: 90
Best Local Similarity: 27.24 Mismatches: 203
Query Match: 14.64 Indels: 156
Db: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-08-894-818B-2 (1-1977)
Qy 62 ValGlySerGluAsnValLysLeuLeu---LysGlyLeuLeuLysLeuGluThrVal 80

Db 103 GTTCAGCAGAAACATTCAGCACTGCTGACCCCGGACCTTCAAGAAAGTCCAGAGATG 162
Qy 81 ProAlaAsnAsnLysLeuHisIleValGlnPheanglyProIleLeuGluThrLys 100
Db 163 AACTGGAACACAGAAAGTGGACACCGTCATTAATGTTCCGG- 201
Qy 101 GlnLysLeuGluThrThrcylValAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
Db 202 -----AGTTAC--- 207
Qy 121 ValGluTyrGluGlyAspValGlnserLysValArgserIleGluHisValGluSerVal 140
Db 208 -----GGAGACAGGACAGGCGCGTTAAGTACGTACAGGCTCAGGCGCCCG 255
Qy 141 GlnProTyrLeuProLysTyrLysIleAspProGlnLeuPheThr-----LysGlyAla 158
Db 256 GTCAAGTAC-----TCCTACAAAGATTAATCCCTGCTGCGCGTTAAATAAAGCCAG 309
Qy 159 SerThrLeuValLysAlaLeuAlaLeuAspThrLysGln-----AsnAsnLysGluVal 176
Db 310 GACCTTCTGCTGATGCGCGGCAATGATACAGCGGTTACTTCGTACCAAGAGGCTTCG 369
Qy 177 GlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIle 196
Db 370 GGCATTAAGTTCATACAGAG----- 390
Qy 197 ThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIle-----ValLys 214
Db 391 -----GATTACAAAGGTTCAAGGTTGACGACGACACTTCCGCTCCCAAGATAGG 438
Qy 215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAla 234
Db 439 GCCGATACCGTGTGAGACTCCCTCGGCTACACAGAGACGGGTGCTGATGCCATCGTC 498
Qy 235 AspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLys 254
Db 499 GATACGGGTATAGACGCAAC-----CACCCGATCTGAAGGCAAG 540
Qy 255 IleThrAlaLeuTyr--AlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHis 273
Db 541 GTCATAGGCTGGTATGACGACCGCTCAACGCGAGTGCAGCCCTACATGACCAAGGACAC 600
Qy 274 GlyThrHisValAlaGlySerValLeuGlyAsnGlyIleThrAsnLys-----Gly 290
Db 601 GGAACCCACGTTGCGGATGCTTGCACGAGCCGACGCGTTACTCCACATGACATAGGC 660
Qy 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
Db 661 GTCCGCCCCGGGCGAAGCTGTCGGCGTCAAGGTTTCGGTCCGACGAGTTCGGGGAAGC 720
Qy 309 LeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAla 328
Db 721 GTCTCCACCATCATGCGCGGTGTGACTGCGTCCAGAACAGAACAGTACCGCGATA 780
Qy 329 ArgIle-----HisThrAsnSerTrpGly 336
Db 781 AGGATCATCAACTTCTCCCTGCGCTCTCCCAAGAGCTCCAGCAACGACCTCCCTCAGT 840
Qy 337 AlaAlaValAsnGlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLys 356
Db 841 CAGGCCGTCAACAAACGCTGGAGCGC----- 867
Qy 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsnGlyTyrThrLysSer 376
Db 868 ---GGTATAGTATGCTGCGTCCGCGCGCAACAGCGGCGCAACACTTACACCTCCGCG 924
Qy 377 AlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSer 396
Db 925 TCACCCGCGCGCGCAGCAAGTCAATACCGTGGTGCA----- 963
Qy 397 PheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThr 416

Db 964 -----GTTGACAGCAAGCAACATCGCAGCTTCTCCAGACGGAGCCAGC 1011
Qy 417 LysAspGlyValArgIleValProAspValMetAlaProGlyThrTyrIleLeuSerAlaArg 436
Db 1012 GCGAGCGAAGGCTGCAAGCGGAGGTGTGCGCCCGCTTGAACATCAATAGCCCGCGC 1071
Qy 437 SerSerLeuAlaProAspSerSerPheTyrAlaAsnHisAspSerIleTyrAlaTyrMet 456
Db 1072 GCCAGC-----GGAACACGATGAGGACCCCGATMAACGACTACACCAAGGCC 1122
Qy 457 GlyGlyThrSerMetAlaThrProIleValAlaGly---AsnValAlaGlnLeuArgGlu 475
Db 1123 TCTGAACACGATGAGCGACCGCGCGCTTGGCGGTGGCGGCTCATCTCCAGGCC 1182
Qy 476 HisPheValIleAsnArgGlyIleThrProLys-----ProSerLeuLeuValAla 493
Db 1183 CAC-----CCGAGCTGACCGCCGACCGCGACGAGGAGAGACCGCC 1218
Qy 494 LeuIleAlaGlyAla-----AlaAspValGlyLeuGlyTyr 505
Db 1219 CTCATGACAGCGCCCGACATAGTCGCCCGCAAGAGATAGCGACATCGCTACGATGCG 1278
Qy 506 ProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSerLeu-----Asn 522
Db 1279 -----GGTAGGCTGAGCTTACAAAGCCATCAAGCTCAATCAAGACGAC 1317
Qy 523 ValAlaTyrValLeuGlnSerSerAlaLeuSerThrSerGlnValAlaThrTyrPhe 542
Db 1318 TAGCGCAAGCTCACTTACCGCGCTCGCGCGACCAAGGAGCGCCACCGACCTTC 1377
Qy 543 ThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaProAlaSer 562
Db 1378 GACGTGACCGCGCCGACCTTGTGACCGCCACCTCTACTGGAGC----- 1422
Qy 563 ThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGly 582
Db 1423 -----ACGGGCTCGAGCGACATCGACTTACCTTACCGACCGCCAAAGGG 1467
Qy 583 ThrArgTyrValGlyLeuAsnAspPheSerAlaProPheAspAsnThrAspGlyAsn 602
Db 1468 AACGAG-----GTTGACTTACTCTTACACCGCTACTAC 1500
Qy 603 AsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGln 622
Db 1501 GCGTTGAGAGAGGTGGCTACTACACCGACCGCGGAGACTGGACGTCAGAGTGTGTC 1560
Qy 623 AlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1561 AGCTACAG-----GGCGGGGCGAAGTACAGGTGACGTCGTCAAGC 1602

RESULT 9
US-09-445-472-11
Sequence 11, Application US/09445472
Patent No. 6358726
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOTO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA=6
CURRENT APPLICATION NUMBER: US/09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 1977
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Synthetic
US-09-445-472-11
Alignment Scores:
Pred. No.: 9,76e-37 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.8% Conservative: 90
Best Local Similarity: 27.2% Mismatches: 203
Query Match: 14.6% Indels: 156
DB: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-09-445-472-11 (1-1977)

Qy 62 ValGluSerGluAsnValIleValLeuLeu---LysGlyLeuLeuLysIleValGluThrVal 80
Db 103 GTTCAGCAGAGAACTACGAGACTGTGACCGCGGAGCTGTTCAGAAAGTCCAGAGATG 162
Qy 81 ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGluGluThrLys 100
Db 163 AACTGAAACCGAAGATGAGACACCTCATTAATGTTGGG----- 201
Qy 101 GlnLysLeuGluIleThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
Db 202 -----AGCTAC--- 207
Qy 121 ValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHisValGluSerVal 140
Db 208 -----GGAACAGGAGCAGGCGCGTTAAGTACTGAGGCTCATGGCGCCAG 255
Qy 141 GluProTyrLeuProLysTyrIleAspProGlnLeuPheThr-----LysGlyAla 158
Db 256 GTCAAGTAC-----TCTTACAGATTAATCCTGCTGTCGCGGTTAAATTAAGCCAGG 309
Qy 159 SerThrLeuValLysAlaLeuAlaLeuAspThrLysGln-----AsnAsnLysGluVal 176
Db 310 GACCTTGTGCTGATCCGCGGACATGATAGACACGGTTACTTCCGTAACACAGGGTCTCG 369
Qy 177 GlnLeuArgGlyIleGluGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIle 196
Db 370 GGCATTAAGTCTTACAGAG----- 390
Qy 197 ThrAlaLysProGluTyrLysValMetAsnAspValAlaArgIle-----ValLys 214
Db 391 -----GATTACAGGTTTCAAGTTGACGAGCCACTTCCTCCGATRAGG 438
Qy 215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnIleValAlaValAla 234
Db 439 GCGGATACCGTGTGAATCTCCGCGCTACGACGAAAGCGGTGTGGTGTGCGATGTC 498
Qy 235 AspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLys 254
Db 499 GATACCGGTATTAACCGCAAC-----CACCCGATCTGAAGGGCGAG 540
Qy 255 IleThrAlaLeuLys---AlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHis 273
Db 541 GTCATAGGCTGTGACAGCGCGTCAACGCGAGGTCAACCCCTTACATGACAGGAGAC 600
Qy 274 GlyThrHisValAlaGlySerValIleGlyAsnGlyAlaThrAsnLys-----Gly 290
Db 601 GGAACCAAGTTCGGGTTATCGTTGCCGGAACCGGAGGTTAATCCCACTACATAGGC 660
Qy 291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGly 308
Db 661 GTGCGCCCGCGCGAAGCTCGCGCGCTCAAGGTCTCTGTCGCCACCGTTTGAGGAGC 720
Qy 309 LeuGlyGlyLeuProSerAsnLeuGlnThrIlePheSerGlnAlaPheSerAlaGlyAla 328
Db 721 GTTCAACCATATCGCGGGGTGTGACTGGGTGTCCAGAAACAGACAAATACGGGATA 780
Qy 329 ArgIle-----HisThrAsnSerTyrGly 336
Db 781 AGGGTATTAACCTTCCCTCGGCTCTCCAGAGCTCCGAGGAAACGACTCTCCAGT 840


```

QY 337 AAlaValaAsnGlyValaThrThrAspSerArgAsnValaAspArgValaArgLys 356
   |||||
   |||||
   |||||
Db 841 CAGGGCGTCAACAACCGCTGGAGCGC----- 867
QY 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGlnArgProAsnGlyThrIleSer 376
   |||||
   |||||
   |||||
Db 868 ---GGTATGATGTCGCGTCCGCGCGCAAGCGGCGCAACCTACACCGTCGCGC 924
QY 377 AlaProGlyThrAlaValaAsnAlaIleThrValaGlyAlaThrGlnAsnLeuArgProSer 396
   |||||
   |||||
   |||||
Db 925 TCACCGCGCGCGCGCAAGCTCATACCGCTGGTGA----- 963
QY 397 PheGlySerThrAlaAspAsnIleAsnIleValaAlaGlnPheSerSerArgGlyProThr 416
   |||||
   |||||
   |||||
Db 964 -----GTTGACAGCAACGACGACATCGCCAGCTTCTCCAGCAGGGAGCGAC 1011
QY 417 LysAspGlyArgGlyLeuArgProAspValaMetAlaProGlyThrThrIleLeuSerAlaArg 436
   |||||
   |||||
   |||||
Db 1012 GCGGACGGAAAGCTCAAGCGCGAAGTCGTGCGCGCGCGGTGACATCATAGCCCGCGC 1071
QY 437 SerSerLeuAlaProAspSerPheThrAlaAsnIleAspSerIleThrAlaThrMet 456
   |||||
   |||||
   |||||
Db 1072 GCGAGC-----GGAACAGCATGGGCAACCCGATAAAGACTACTACCAAGCGC 1122
QY 457 GlyGlyThrSerMetAlaThrProIleValaAlaGly---AsnValaAlaGlnLeuArgGln 475
   |||||
   |||||
   |||||
Db 1123 TCTGGAACGAGATGGCCACCGCGACGTTTGGGGGTGGCGGCTACTCTCCAGCGC 1182
QY 476 HisPheValaLysAsnArgGlyIleThrProLys-----ProSerLeuLeuLysAlaAla 493
   |||||
   |||||
   |||||
Db 1183 CAC-----CCGAGCTGGACCCCGGACACAGAGGTGAAGACCGCGC 1218
QY 494 LeuIleAlaGlyAla-----AlaAspValaGlyLeuGlyThr 505
   |||||
   |||||
   |||||
Db 1219 CTCATACGAGACCGCGCACTAGTCCGCCCAAGAGATGCGGACATCGCCCTACGCTGGC 1278
QY 506 ProAsnGlyAsnGlnGlyTrpGlyArgValaThrLeuAspLysSerLeu-----Asn 522
   |||||
   |||||
   |||||
Db 1279 -----GCTAGCGTGAACGCTTACAGGCGATCAAGTACAGACGAC 1317
QY 523 ValAlaThrValaAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrThrPhe 542
   |||||
   |||||
   |||||
Db 1318 TACGCGCAACTCACCCTTACCGGCTCCGCGCGCAAGGAAAGCGCCACCCACCTTC 1377
QY 543 ThrAlaThrAlaGlyLysProLeuLysIleSerLeuValaTrpSerAspAlaProAlaSer 562
   |||||
   |||||
   |||||
Db 1378 GACGTCAGCGGCGCCACCTTGTGACCGCCACCTCTACTGGGAC----- 1422
QY 563 ThrThrAlaSerValaThrLeuValaAsnAspLeuAspLeuValaIleThrAlaProAsnGly 582
   |||||
   |||||
   |||||
Db 1423 -----ACGGGCTCGAGCGACATCGACCTTACTCTACGACCCCAAGCGG 1467
QY 583 ThrArgThrValaGlyAsnAspPheSerAlaProPheAspAsnAlaTrpAspGlyArgAsn 602
   |||||
   |||||
   |||||
Db 1468 AACGAG-----GTTGACTACTCTACACCCGCTACTAC 1500
QY 603 AsnValaGluAsnValaPheIleAsnSerProGlnSerGlyThrThrIleGluValaGln 622
   |||||
   |||||
   |||||
Db 1501 GCGTTCGAGAGAGTGGCTACTACACCGGACCGGAACTGGAGCGGTCAAGGTCGTC 1560
QY 623 AlaThrAsnValaProValaGlyProGlnAsnPheSerLeuAlaIleValaAsn 639
   |||||
   |||||
   |||||
Db 1561 AGCTAFAAG-----GGCGCGGCAACTACAGGTGACGTGTGACG 1602

```

RESULT 10

```

US-10-090-624-11
; Sequence 11, Application US/10090624
; Patent No. 6783970
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMODO, Tomoko
; APPLICANT: ASADA, Kiyozo

```

```

; APPLICANT: KATO, Ikunobu
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 11
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; US-10-090-624-11

```

Alignment Scores:

Pred. No.:	9,766-37	Length:	1977
Score:	477.00	Matches:	168
Percent Similarity:	41.84	Conservative:	90
Best Local Similarity:	27.24	Mismatches:	203
Query Match:	14.64	Indels:	156
DB:	3	Gaps:	25

US-10-784-870-4 (1-639) x US-10-090-624-11 (1-1977)

```

QY 62 ValGluSerGluAsnValaLysLeuLeu---LysGlyLeuLeuLysLeuGluThrVal 80
   |||||
   |||||
   |||||
Db 103 GTTCAGCAGAAAGACTACGAGCTGCTGACCCCGGAGCTGTCAAGAAAGTCCAGAGATG 162
QY 81 ProAlaAsnLysLysLeuHisIleValaGlnPheAsnGlyProIleLeuGluThrLys 100
   |||||
   |||||
   |||||
Db 163 AACTGGAAACGAGAGTGAACACCGCTCAATGTTCCGG----- 201
QY 101 GlnLysLeuGluThrThrGlyAlaLysIleLeuAspArgIleProAspArgAlaThrIle 120
   |||||
   |||||
   |||||
Db 202 -----AGCTAC--- 207
QY 121 ValGluThrGluGlyAspValaGlnSerLysValaArgSerIleGluHisValaGluSerVal 140
   |||||
   |||||
   |||||
Db 208 -----GGAGACAGGAGACAGCGCGGTTAAAGTACGAGGCTGATGGCGCCAG 255
QY 141 GluProGlyLeuProLysThrGlyLysIleAspProGlnLeuPheThr-----LysGlyAla 158
   |||||
   |||||
   |||||
Db 256 GTCAAGTAC-----TCTTACAAAGTAAATCCCTGCTGCGCGGTTAAATAAAGGCCAG 309
QY 159 SerThrLeuValaLysAlaLeuAlaLeuAspThrLysGln-----AsnAsnLysGluVal 176
   |||||
   |||||
   |||||
Db 310 GACCTTCTGCTGATCGCGGCGATGATACACGGGGTTACTTCGGTTACACAAAGGTCGCG 369
QY 177 GlnLeuArgGlyIleGluGluIleAlaGlnThrValaAlaSerAsnAspValaHisThrIle 196
   |||||
   |||||
   |||||
Db 370 GGCATAAAGTTCATACAGAG----- 390
QY 197 ThrAlaLysProGluThrLysValaMetAsnAspValaAlaArgGlyIle-----ValLys 214
   |||||
   |||||
   |||||
Db 391 -----GATTCAAGAGTTCAAGTTACACACGCTTCGCTCCAGATAGGG 438
QY 215 AlaAspValaAlaGlnSerSerThrGlyLeuThrGlyGlnGlyGlnIleValaAlaValaAla 234
   |||||
   |||||
   |||||
Db 439 GCCATACCGCTGTGAACCTCCCTCGGCTACGACGAAAGCGGTGTGGTGGTCCATGCTC 498
QY 235 AspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLys 254
   |||||
   |||||
   |||||
Db 499 GATACGGGTAAAGACGGGAC-----CACCCGATCTGAAGGGCAG 540
QY 255 IleThrAlaLeuThr---AlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHis 273
   |||||
   |||||
   |||||
Db 541 GTCATAGGCTGTGACGAGCGCGTCAAGCGCAGGTGACCCCTTACGATGACCAAGGACAC 600

```



```

QY 274 G1YThrh1aVala1aG1ySerVala1eug1yAaNg1yAlaThraAa1yS-----G1Y 290
DB 601 GGAACCCACGCTGGCGGATGCTGGCCGGAACCGGACCGTTAACTCCACATACATAGGC 660
QY 291 MetAlaProGlnAlaAa1eVala1PheGlnSer1IeMet-----AaPseSerG1yG1Y 308
DB 661 GTGCCCCCGCGGGAAGCTCGCGGCTCAAGGTTCTCGTGCCGACGCTTCGGGAAGC 720
QY 309 LeuG1yG1LeuProSerAa1eug1nThrLeuPheSerG1nAlaPheSerAlaG1yAla 328
DB 721 GTCTCCACATCATCGCGGGGTGTGACTGGGTCTCCAGAACAGACCAATACGGAATA 780
QY 329 Arg1Ie-----HisThraAaSerTrpG1y 336
DB 781 AGGCTCATCAACCTCTCCCTGGCTCTCCAGAGCTCCAGCGAACCACTCCCTCACT 840
QY 337 AlaAlaVala1aNg1yAla1yTrhThraAaSerArgAaVala1aAaP1yTrVala1yA 356
DB 841 CAGGCTCTCAACCAACGCTGGGACGCC----- 867
QY 357 AaAaPheThrl1eLeuPheAla1aG1yAaNg1yAa1yAa1yAa1yAa1yAa1yAa1y 376
DB 868 ---GGTATAGTAGTCTGCTGCTGCGCGGCAACGCGGCGGACCTACACCGTGGC 924
QY 377 AlaProG1yThra1a1yAaAa1a1eThraValG1yAlaThraG1uAa1eAa1yAa 396
DB 925 TCACCGCGCGCGGACCAAGTCAACCTCGGTCA----- 963
QY 397 PheG1ySerTrAlaAaPaa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1e 416
DB 964 -----GTTGACACCAACCAACATCCGCTTCACAGGAGGACGAC 1011
QY 417 LysAaP1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1y 436
DB 1012 GCGAGCAAGGCTCAAGCGGAAGTGTGCGCGCGCTTGAATCAATCAATCAATCAAT 1071
QY 437 SerS1rLeuAlaProAaPseSerPheTrAlaAa1eAaPseSer1yTrAla1yMet 456
DB 1072 GCCAGC-----GGAACCAAGATGGGACCCCGATTAACATCACTACCAAGGCC 1122
QY 457 G1yG1yThraSerMetAlaThraPro1eVala1aG1y---AaVala1aG1eAa1y 475
DB 1123 TCTGGAACCAAGTCAAGCGGACCGGACGTTGGCGGCTCATCTCCAGGCC 1182
QY 476 HisPheValyAaAaNg1y1eThraPro1yS-----ProSerLeuLeu1yAla1a 493
DB 1183 CAC-----CCGAGCTGACCCCGGACCAAGGAGGAGACCGCC 1218
QY 494 Leu1eAlaG1yAla-----AlaAaPValG1yLeuG1yTr 505
DB 1219 CTATCGAAGACCGCCGACATAGTCGCCCCCAAGAGATAGCGGACATCGCTACGGTGC 1278
QY 506 ProAaNg1yAaNg1yTrpG1yArgVala1ThraAaP1ySerLeu-----Aa 522
DB 1279 -----GGTAGGTGACGTCAAGGCAAGGCAATCAAGTACAGAC 1317
QY 523 ValAla1yTrVala1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1e 542
DB 1318 TAGCGAAGCTCACTTCACCGGCTCGTCCGACCAAGGAGGACCGCACCTCC 1377
QY 543 ThrAlaThra1aG1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1y 562
DB 1378 GACGTCAAGCGCGCACCTTCGTAACCGCACTCTACTGAGAC----- 1422
QY 563 ThrThra1aSerValThraLeuValAaAaP1eAaP1eVal11eThra1aProAaNg1y 582
DB 1423 -----ACGGGCTCGAGGACATGACCTCACTCACTCAAGCCCAACGCG 1467
QY 583 ThrArgTrVala1yAaAaPheSerAlaProPheAaPaa1eAa1yAaAa 602
DB 1468 AACGAG-----GTTGACTACTCTCAACCGGCTACTAC 1500
QY 603 AaVala1G1uAaVala1Phe1eAaAaSerProGlnSerG1yThra1yThrl1eG1uValGln 622

```

```

DB 1501 GGCTTCAGAAAGCTCGGCTACTACACCGACCGCGGACCTCGACGCTCAAGTCTC 1560
QY 623 AlaTrAaAaValProValG1yProGlnAaAaPheSerLeuAla1eVala1a 639
DB 1561 AGCTAACAG-----GGCGCGCGGACACTACACGCTCAAGTCTCAAGTCTCAGC 1602

RESULT 11
US-09-841-553-2
; Sequence 2, Application US/09841553
; Patent No. 6849441
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; MORISHITA, Mio
; YAMAMOTO, Katsuhiko
; MITTA, Masanori
; ASADA, Kiyozo
; TSUNASAWA, Suenmu
; KATO, Ikunobu
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; City: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,553
; FILING DATE: 24-Apr-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/894,818
; FILING DATE: <Unknown>
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: TAKAKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-841-553-2

Alignment Scores:
Pred. No.: 9,76e-37 Length: 1977
Score: 477.00 Matches: 168
Percent Similarity: 41.8% Conservative: 90
Best Local Similarity: 27.2% Mismatches: 203
Query Match: 14.6% Indels: 156
DB: 3 Gaps: 25

US-10-784-870-4 (1-639) x US-09-841-553-2 (1-1977)
QY 62 ValG1uSerG1uAaAaVal1yAa1eAa1yAa1yAa1yAa1yAa1yAa1yAa1yAa1y 80
DB 103 GTTCAGACAGAAAGTCAAGGACCTGTAACCGGACGCTTCAAGAAAGTCAAGGAGT 162
QY 81 ProAlaAaAaAa1yAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1eAa1e 100

```



```

Db 964 -----GTTGACAAAGTATGATGTTATTAACAAGCTTCTCA 996
Qy 412 SerATGGLYProThrLYAspGLYArgILEYserProAspVALMetALAProGLYThrTYR 431
Db 997 ACCAGAGGGCCAACTGACAGACGGCTTAAGCTGAGGTGGTCTCAAGAACTGG 1056
Qy 432 ILEUSeRALArgSerSerLEUVALProAspSerSerPheTRPAlaAsnHISAspSer 451
Db 1057 ATPATGTGCTCCAGAGCAAGT-----GGAACTAGCATGGGTCAACCAATTATATGAC 1107
Qy 452 LysTYRAlaTYRMeGLYGLYThrSerMetALAThrProILEVALAlaGLYAsnVALAla 471
Db 1108 TATTACACAGACAGCTCTCGGACATCAATGGCAACTCTCAAGTACCTGGATTGGACGC 1167
Qy 472 GlnLEUArgGLNHisPheVALysAsnArgGLYILEThrProLYS-----ProSerLEU 489
Db 1168 CTCTTCTCCAA-----GCACACCCGAGCTGAGCTCCAGACAA 1206
Qy 490 LeuLYAlaAlaLeuILEAlaGLYAla-----AlaAspVAL 501
Db 1207 GTAAAAAAGCCCTCATTAAGAACTGCTGATATGTAAGCAGATGAAATAGCCGATATA 1266
Qy 502 GlyLEUGLYTYRProAsnGLYAsnGLNGLYTRPGLYArgVALThrLEUAspLYSerLEU 521
Db 1267 GCCTACGGTGA-----GTTAGGGTTATATGATCAAGAGCTATA 1305
Qy 522 AsnVALAlaTYRValAsnGLYSerSeraLEUSeRThSerGLNLYSAla----- 538
Db 1306 AAC-----TACATATACATATGCAAGCTGCTGATCAAGATATGTCACAAAGGC 1359
Qy 539 -----ThrTYRThrPheThrAlaThrAlaGLYLYSProLEUValILESerLEUValITRP 556
Db 1360 AGCCAAACTCAGACAGTTCGTTATTAAGCGAGCTTCGTTCTTAAGTCCCATTAATACG 1419
Qy 557 SerAspAlaProAlaSerThrThrAlaSerVALThrLEUValAsnAspLEUAspLEUVal 576
Db 1420 GACAAATGCCAAT-----AGCACTTGAATCTTTAC 1449
Qy 577 ILEThrAlaProAsnGLYThrArgTYRValGLYAsnAspPheSeraLProPheAspAsn 596
Db 1450 CTCTACGATCCCAATGGAACAG-----GTTGACTAC 1482
Qy 597 AsnTRPAspGLYArgAsnAsnVALGLYAsnVALPheILEAsnSerProGLNAsnSerGLYThr 616
Db 1483 TCTTACACCCGCTACTATGATTGCAAAAGTGGTATTATTAACCACTGATGAAACA 1542
Qy 617 TYRThrILEGLYValGLNAlaTYRAsnVALProVALGLYProGLNAsnPheSerLEUAla 636
Db 1543 TGGCAATTAAAGTTTGAAGTACAGC-----GGAAGTGAACATCAAGTATGAT 1593
Qy 637 ILEValAsn 639
Db 1594 GTGGTAAGT 1602

```

```

STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,553
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/894,818
FILING DATE: <unknown>
APPLICATION NUMBER: JP 323285/1995
FILING DATE: 12-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1962 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-841-553-34
Alignment Scores:
Pred. No.: 2,116-33 Length: 1962
Score: 443.00 Matches: 163
Percent Similarity: 40.3% Conservative: 88
Best Local Similarity: 26.2% Mismatches: 204
Query Match: 13.6% Indels: 168
Gaps: 25
US-10-784-870-4 (1-639) x US-09-841-553-34 (1-1962)
Qy 62 VALGLYSerGLYAsnVALLYSLeuLEUValLYSLeuGLYValLYSLeuGLYThrVALPro 81
Db 103 GTTGAGAAACTATAGTGTCTTAAGCCAGGACTGTTCAAGAAATTCAAAAATTCGAAAT 162
Qy 82 ALAAsnAsnLYSLeuHISILEVALGlnPheAsnGLYProILEUGLNGLNThrTYRGLN 101
Db 163 CCTTAACGAGAAATCAGACAGTA-----ATTGTAATTGAAAAACCATAGGGGA 210
Qy 102 Lys-----LEUThrThrGLYAlaLYSILEUAspTYRILEPro 115
Db 211 AAAAATTGCACTAAGAGTCTTGAATTAAAGGTGCAAAAGTT----- 255
Qy 116 AspTYRAlaTYRILEVALGlnTYRGLNGLYAspVALGlnSerLYSValArgSerILEGLN 135
Db 256 -----AGGTATGTC----- 264
Qy 136 HISValGLYSerVALGlnProTYRLEUProLYSILEAspProGLNLeuPheThr 155
Db 265 -----TACCATATTATACCCCAATA----- 285
Qy 156 LysGLYAlaSerThrLEUValLYSAlaLEUAlaAspThrLYSGLNAsnAsnLYSGLN 175
Db 286 ---GCTGCCGATCTTAAGTTAGAGCTTACATGATCATGAGTTTAAACAGGGGCTAAA 342
Qy 176 VALGlnLEUArgGLYILEGLNGLNILEAGlnTYRVALAlaSerAsnAspVALHISTYR 195
Db 343 GCTAACCTTTCAGGTGTAGGTTATC-----CAGGAAGACTCAAA 384
Qy 196 ILEThrAlaLYSProGLNLYRYSValMetAsnAspVALAlaArgGLYILEVALYVal 215

```

RESULT 15
 US-09-841-553-34
 Sequence 34, Application US/09841553
 Patent No. 6849441
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 MORISHITA, Mio
 YAMAMOTO, Katsuhiko
 MITTA, Masanori
 ASADA, Kiyozo
 TSUNASAWA, Susumu
 KATO, Kunobih
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESS: Browdy and Neimark
 STREET: 419 Seventh Street N.W., Ste. 300
 CITY: Washington

Db 385 GTTACAGTTTCAGCAGAACTAGAGACTGTCGAGCTCAAGTTATGGCACT 444
 Qy 216 AspValAlaGlnSerSerThrGlyLeuThrGlyGlnGlnIleValAlaValAlaAsp 235
 Db 445 TACGTTTGAACTTGGGATATGAT-----GGTCTGGAATTCACATAGAAATTAATTAC 498
 Qy 236 ThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLeuIle 255
 Db 499 ACTGGAAATTCAC-----GCTTCTCATCCAGATCTCCAGAGAAAGTA 540
 Qy 256 ThrAlaLeuThrAlaLeuGlyArgThrAsnAsnAlaAsn-----Asp 269
 Db 541 -----ATTGGGTGGGTAGATTCTTTCATGAGAGATTATTCATACAT 585
 Qy 270 ThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGlyAsnGlyAlaThrAsn--- 288
 Db 586 GACCATGACATGAACTCATGATGATTCATATAGACGCTGTATCTGGAGCAGCAAGTAT 645
 Qy 289 -----LysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMet----- 303
 Db 646 GCGAAGTACAAAGGAATGCTCCAGAGCTAAGCTGGCGGAATTAAGTTCTAGTCC 705
 Qy 304 AspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAla 323
 Db 706 GATGTTCTGGAGACATATCTACTATATTAAGGAGTTGAGTGGCCGTTGATTAACA 765
 Qy 324 PheSerAlaGlyAlaArgIleHisThrAsnSerThrGlyAla----- 337
 Db 766 GATTAAGTACGGAATTAAGTATTAATCTTCTCTGTTGTCAGCAGAGCTCAGATGCT 825
 Qy 338 -----AlaValAsnGlyAlaThrThrAspSerArgAsnValAsp 351
 Db 826 ACTGACGCTCTAAGCTCAGGCTGTTAATGCAAGCTGGAGTCT----- 867
 Qy 352 AspThrValArgGlyAsnAspMetThrIleLeuPheAlaAlaGlyAsnGluArgProAsn 371
 Db 868 -----GATTAAGTTGTTGTTGTTGCTGGGAAACAGTGAACCTTAC 909
 Qy 372 GlyGlyThrIleSerAlaProGlyThrAlaLeuAsnAlaIleThrValGlyAlaThrGlu 391
 Db 910 AAGTATCAATCGGTTCTCCAGCAGCTGCAGCAAAAGTTATTACAGTTGGAGCC----- 963
 Qy 392 AsnLeuArgProSerPheGlySerThrAlaPheAsnIleAsnIleValAlaGlnPheSer 411
 Db 964 -----GTTGACAAAGTATGATGTTTATTAACAAGCTTCTCA 996
 Qy 412 SerArgGlyProThrIleAspGlyArgIleLeuProAspValMetAlaProGlyThrTyr 431
 Db 997 AGCAGAGGCGCAACTGCAAGCGGCTTAAGCTTGAAGTTGTTGCTCCAGGAAACTGG 1056
 Qy 432 IleLeuSerAlaArgSerSerLeuAlaProAspSerPheThrAlaAsnHisAspSer 451
 Db 1057 ATTAATTCCTGCAGAGCAAGT-----GGAACCTAGCATGGGTCAACCAATTAATGAC 1107
 Qy 452 LysThrAlaThrMetGlyThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
 Db 1108 TATTACACAGAGCTCTGAGACATCAATGCAACTCTCACTGAGCTGATTGACGCC 1167
 Qy 472 GlnLeuArgGlnHisPheValIleAsnArgGlyIleThrProLys-----ProSerLeu 489
 Db 1168 CTCTGTCTCCA-----GCACACCCGAGCTGAGCTCCAGACAA 1206
 Qy 490 LeuLeuAlaAlaLeuIleAlaGlyAla-----AlaAspVal 501
 Db 1207 GTAAACACAGCCCTCATAGAACTGCTGATATCTAAAGCCAGATGAATAGCCGATATA 1266
 Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyThrGlyArgValThrLeuAspLysSerLeu 521
 Db 1267 GCTTACGCTGCA-----GTTAGGTTAATGCAATCAAGGCTATA 1305
 Qy 522 AsnValAlaThrValAsnGlnSerSerAlaLeuSerThrSerGlnLysAla----- 538
 Db 1306 AAC-----TACGATTAATGCAAAAGCTAGTGTCTGATATGTTGCCAACAAGCC 1359

Qy 539 -----ThrThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTrp 556
 Db 1360 AGCCAACTTCCACAGTTCTGTTATTAGCGAGCTTCTGTTACTGTCACATTAATCTGG 1419
 Qy 557 SerAspAlaProAlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuVal 576
 Db 1420 GACATATCCAT-----ACGACCTTGAATCTTAC 1449
 Qy 577 IleThrAlaProAsnGlyThrArgThrValGlyAsnAspPheSerAlaProPheAspAsn 596
 Db 1450 CTTCACATCCCAATGGAACCAAG-----GTTGACTAC 1482
 Qy 597 AsnThrAspGlyArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThr 616
 Db 1483 TCTTACACCGCTACTATGATTCGAAAGGTTGTTATTACAAACCAACTGATGAAACA 1542
 Qy 617 TyrThrIleGluValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAla 636
 Db 1543 TGGCAATTTAAGGTTGTAAGCTACAGC-----GAAAGTCAAACTATCAAGTATGAT 1593
 Qy 637 IleValAsn 639
 Db 1594 GTGTTAAGT 1602

Search completed: April 8, 2006, 03:53:40
 Job time : 585 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: April 8, 2006, 03:55:21 ; Search time 1318 Seconds

(without alignments)
4009.205 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267
Sequence: 1 MRKKVFLSLTSLAAALISTV.....EVQAVNYVPGQNPSTLAIVN 639

Scoring table:

BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/abs/ABSSWB.spool/US10784870/runat_07042006_143427_27724/app_query.fasta_1
-DB=Published Applications NA Main -GFWT=fastap -SUFFIX=p2n.rnpbm
-MIMATCH=0.1 -LOOPEC=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1
-MATRIX=blonum62 -TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext
-HEP512=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs804
-USRR=US10784870 @CCN 1.1 1026 @runat_07042006_143427_27724 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3267	100.0	1920	3	US-09-920-954-3
2	3267	100.0	1920	7	US-10-784-870-3
3	3072	94.0	1923	3	US-09-920-954-7
4	3072	94.0	1923	7	US-10-784-870-7
5	3066	93.8	1923	3	US-09-920-954-5
6	3066	93.8	1923	7	US-10-456-479-3
7	3066	93.8	1923	7	US-10-784-870-5

8	3066	93.8	1923	8	US-10-820-712A-2	Sequence 2, Appl1
9	3066	93.8	1923	9	US-10-820-714A-2	Sequence 2, Appl1
10	2237	68.5	1302	8	US-10-820-712A-11	Sequence 11, Appl1
11	2237	68.5	1302	9	US-10-820-714A-12	Sequence 12, Appl1
12	2188	67.0	1302	8	US-10-820-712A-13	Sequence 13, Appl1
13	2188	67.0	1302	9	US-10-820-714A-14	Sequence 14, Appl1
14	2183	66.8	1305	6	US-10-385-662-1	Sequence 1, Appl1
15	2122	65.0	1302	8	US-10-820-712A-21	Sequence 1, Appl1
16	2122	65.0	1302	9	US-10-820-714A-22	Sequence 22, Appl1
17	2002.5	61.3	1299	8	US-10-820-712A-19	Sequence 19, Appl1
18	2002.5	61.3	1299	9	US-10-820-714A-20	Sequence 20, Appl1
19	1998.5	61.2	1299	8	US-10-820-712A-15	Sequence 15, Appl1
20	1998.5	61.2	1299	9	US-10-820-714A-16	Sequence 16, Appl1
21	1991.5	61.0	1299	8	US-10-820-712A-17	Sequence 17, Appl1
22	1991.5	61.0	1299	9	US-10-820-714A-18	Sequence 18, Appl1
23	477	14.6	1977	3	US-09-841-553-2	Sequence 2, Appl1
24	477	14.6	1977	5	US-10-090-624-11	Sequence 11, Appl1
25	477	14.6	1977	8	US-10-800-684-2	Sequence 2, Appl1
26	477	14.6	1977	9	US-10-888-588-11	Sequence 11, Appl1
27	443	13.6	1962	3	US-09-841-553-14	Sequence 34, Appl1
28	443	13.6	1962	5	US-10-090-624-15	Sequence 15, Appl1
29	443	13.6	1962	8	US-10-800-684-14	Sequence 34, Appl1
30	443	13.6	1962	9	US-10-888-588-15	Sequence 15, Appl1
31	432	13.2	1977	3	US-09-841-553-6	Sequence 6, Appl1
32	432	13.2	1977	8	US-10-800-684-6	Sequence 6, Appl1
33	417.5	12.8	1236	5	US-10-090-624-2	Sequence 2, Appl1
34	417.5	12.8	1236	9	US-10-888-588-2	Sequence 2, Appl1
35	417.5	12.8	1566	3	US-09-841-553-4	Sequence 4, Appl1
36	417.5	12.8	1566	8	US-10-800-684-4	Sequence 4, Appl1
37	359	11.0	3624	6	US-10-156-761-5701	Sequence 5701, Ap
38	359	11.0	9025608	6	US-10-156-761-1	Sequence 1, Appl1
39	357.5	10.9	3417	5	US-10-156-761-3306	Sequence 3306, Ap
40	357.5	10.9	9025608	6	US-10-156-761-1	Sequence 1, Appl1
41	344.5	10.5	135638	6	US-10-314-657-1	Sequence 1, Appl1
42	344.5	10.5	135638	9	US-10-473-193-1	Sequence 1, Appl1
43	342.5	10.5	1339	3	US-09-974-300-1934	Sequence 1934, Ap
44	326	10.0	4765	3	US-09-841-553-7	Sequence 7, Appl1
45	326	10.0	4765	5	US-10-090-624-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-09-920-954-3
Sequence 3, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUMI
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURES:
NAME/KEY: CDS

```

LOCATION: (1) .. (1920)
US-09-920-954-3

Alignment Scores:
Pred. No.: 1,226-287      Length: 1920
Score: 3267.00           Matches: 639
Percent Similarity: 100.0%   Conservative: 0
Best Local Similarity: 100.0%   Mismatches: 0
Query Match: 100.0%         Indels: 0
DB: 3                      Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-3 (1-1920)

QY 1 MetArgbysbysbysValPheLeuSerValLeuSerAlaAlaAlaIleuSerThrVal 20
DB 1 ATGAGAAAGAAAGAGGCTTTTATCTGTTTATCAGCTGACGAGATTCGTGACGCT 60
QY 21 AlaLeuAnaAnPProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLySGLyIle 40
DB 61 GCATTAAACAATCCCTCGGCTGGTGATGCAAGGACTTTGATCTGGATTTTAAAGGAT 120
QY 41 GluThrThrThrAspValSerGlyPheSerLyGlnArgGlnThrGlyAlaAlaAlaPhe 60
DB 121 CAACAACAACAACGATGTCAGTGTCTTCCAAACAAGCAACAACAGGTGCGCTGCAATT 180
QY 61 LeuValGluSerGluAsnValLysLeuLeuLyGlyLeuLeuLysbysbysLeuGluThrVal 80
DB 181 CTGGTGAAGCTGAAAATGTGAACCTTTAAAGATGCTTAAAGAACTTGAACAAGTA 240
QY 81 ProAlaAnaAnLysLeuHsIleValGlnPheAsnGlyProIleLeuGluGluThrLys 100
DB 241 CCGGCAAAATATTAACCTCATATTTGTCATATGCAATGCCCATTTTAAAGAAACAATA 300
QY 101 GluLysLeuGluThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaIle 120
DB 301 CAAGAAGTAAAGACAATGAGCAAGATTCGACATCAATCCCTGATTAATGCAATATAT 360
QY 121 ValGluTyrGluGlyAspValGlnSerLysValArgSerIleGluHsIleValGluSerVal 140
DB 361 GTGAGATATGAGGGGGATGTTTCACTGCAAAAGTCGCTCCATTGAACAAGTGAATAGTG 420
QY 141 GluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
DB 421 GAGCCCTTACTTGCAGAAATCAAAATAGATCCCACTTTTCAAAAAGCGCATGACG 480
QY 161 LeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgGly 180
DB 481 CTGGTAAAGCGTTGGCGCTTGATACAGACAGAAACAATTAAGATGCAATTAAGAGGC 540
QY 181 IleGluGluIleAlaGlnTyrValAlaSerAsnAspValHsIleTyrIleThrAlaLysPro 200
DB 541 ATCGAAGGAATCGCTCAGTACGTAGCAAGCAATGACATTAATTAATTCGGCAAGCCT 600
QY 201 GluTyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
DB 601 GAATATAAGGTGATGATGATGATGCGCAAGGATATGTCAAAGCGATGTGGCAAGAC 660
QY 221 SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrLysLeuAspThr 240
DB 661 AGCTACGCTTTGATGAGACAGGCGCAAGATGTGCGAGTTGCGATCTGAAATGGATTA 720
QY 241 GlyTyrAsnAspSerSerMetHsIleGluAlaPheArgGlyLysIleThrAlaLeuTyrAla 260
DB 721 GGAAGAAACGACAGTTCGATGATGAGAGCTTCGGCGTAAATAACACACATTAATGCA 780
QY 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHsIleGlyThrHsIleValAlaGlySer 280
DB 781 CTGGGTCCGACAAATATATGCAATGATGCAAGACGGTCAATGATACCATGTGGCAAGTTCG 840
QY 281 ValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
DB 841 GATTATGAAATGGCGCAACGAATTAAGAAATGGCACTCCAGCGCAATCTGGATTTTCA 900

```

```

QY 301 SerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe 320
DB 901 TCCATCATGATATAGAGTGAGTGGCTTGGAGCTTGCATTCGAAATCTTAATTC 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHsIleThrAsnSerTyrGlyAlaAlaValAsn 340
DB 961 AGCCAAAGCATTCAGTGCAGAGTGCAGAAATTCATCAAAATCTCTGGGGGGCAGCGGTGAAT 1020
QY 341 GlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThr 360
DB 1021 GGGGCTTACACACAGATTCAGAAATGTGATGATCTATGTAAGGAAATGATATGACG 1080
QY 361 IleLeuPheAlaAlaGlyAsnGlyLysArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
DB 1081 ATTCCTTTGCGAGTGGGAATAAAGCGCGGAACGCGGTACATGACATGCACTGGATACG 1140
QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyr 400
DB 1141 GCTTAAAAAAGCATTAACAGTCCGCGCAACCGAAACCTGCGTCCAAAGCTTCGGTTCCTAT 1200
QY 401 AlaAspAsnIleAsnHsIleValAlaGlnPheSerSerArgGlyProThrLysAspGlyArg 420
DB 1201 GCAGATTAATATTAACACAGTTCACAGATTCCTCCGCGCGCAAAAGATGGCGCA 1260
QY 421 IleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAAGCTGATGTCATGGCGCCAGGAGACATTAATTAATGCAAAAGATCTCTTGCA 1320
QY 441 ProAspSerSerPheTyrAlaAsnHsIleAspSerLysTyrAlaLysMetGlyGlyThrSer 460
DB 1321 CCGGATTCCTCTTCTGGGCGAATCATGACAGCAAAATATGCTTAATGGGTGAACGTCC 1380
QY 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuAspGluHsIlePheValLysAsn 480
DB 1381 ATGGCAACAACGATATGTGGCGGAAATGTGACAGCTCCGAGCATTTTGTGAATAAT 1440
QY 481 ArgGlyIleThrProLysProSerLeuLysValAlaLeuIleAlaGlyAlaAlaAsp 500
DB 1441 AGAGAAATCACTCTTAAGCTTCTCTTAATGAAGCAAGCTTGAATGGAGTGCGCTGAT 1500
QY 501 ValGlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSer 520
DB 1501 GTTGATATGGGTATCCGAACGAAACCAAGATGGGCGAGTGAACCTGGATTAATCG 1560
QY 521 LeuAsnValAlaTyrValAsnGlyLeuSerSerAlaLeuSerThrSerGlnLysAlaThrTyr 540
DB 1561 TTGAACGTGGCTTAATGTAAGAAATCCAGTCCCTATCACTAGCCAAAGCGACATAT 1620
QY 541 ThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValTyrSerAspAlaPro 560
DB 1621 AACTTTACTGCAACGCGCGGCAAGCATTTGAATAATCTCCCTGGTATGGTCCGATGCCCT 1680
QY 561 AlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
DB 1681 GCAAGCACTACTGCTTGTGTAACCTGGGTCAATGATTTGGATTGTCATTACAGCAACA 1740
QY 581 AsnGlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTyrAspGly 600
DB 1741 AACGGAACAAGATATGTGCGGAATGACTTCCACACACATTTGCAATTAACCTGGAGTGGC 1800
QY 601 ArgAsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyTyrTyrThrIleGlu 620
DB 1801 CGCAATTAACGTAGAAATATGATTTTATTAATTCGCCCAAAATGGAACATATACATTAAG 1860
QY 621 ValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1861 GTGCAAGCATATATATGTGCGCGTGGACCACAAAATCTCTGTTGGCAATGTGAAC 1917

```

RESULT 2
 US-10-784-870-3
 ; Sequence 3, Application US/10784870
 ; Publication No. US20040142837A1
 ; GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKURA, SHITSUMI
APPLICANT: NOMURA, MASATSUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784, 870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509, 814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-10-784-870-3

Alignment Scores:

Pred. No.:	1,22e-287	Length:	1920
Score:	3267.00	Matches:	639
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-784-870-4 (1-639) x US-10-784-870-3 (1-1920)

QY 1 MetArgLysAlaValPheLeuSerValLeuSerAlaAlaAlaLeuSerThrVal 20
DB 1 ATAGAAAGAAAGAGGTGTTTATCTGTTTATCAGCTCAGCGATTCTGCGACTGTT 60
QY 21 AlaLeuAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40
DB 61 GCATTAAACATCCCTCGCTGCGATGCAAGACATTGATCGAATTTAAAGGAATT 120
QY 41 GlnThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaPhe 60
DB 121 CAACCAACACCATGTCAGTGGTTTCTCCAAACAGGACAAACAGGTGGCGCTGCAATT 180
QY 61 LeuValGlnSerGlnAsnValLysLeuLeuLysGlyLeuLeuLysLysLeuGlnThrVal 80
DB 181 CTGGTGAAGCTGAAAATGTGAACCTTCTTAAGGATTCCTAAAGAACTTGAACAGTA 240
QY 81 ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGlnThrLys 100
DB 241 CCGGCAAAATATAACCTCCATATGTCATTCATTCAGTCCCATTTTGAAGAAACAAA 300
QY 101 GlnLysLeuGlnThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaLysTyrIle 120
DB 301 CAGACGTAAGACCACTGAGCAAGCAAGATTCTGCACTACATCCCTGATATGCAATATTT 360
QY 121 ValGlnTyrGlnGlyAspValGlnSerLysValArgSerIleGlnHisValGlnSerVal 140
DB 361 GTGAGATATGAGGGGAGTTCAGTCAAAAGTCGCTCCATTGAACAAGTGAATCAGTG 420
QY 141 GlnProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
DB 421 GAACCACTACTTGCAGAAATCAAAATGATCCCAAGCTTTTCAAAAAGCGATCAACG 480
QY 161 LeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlnValGlnLeuArgGly 180

DB 481 CTGGTGAAGCGTTGGCGCTTGATACGAGCAGAACATAAAGATGCAATTAAGAGCG 540
QY 181 IleGlnGluIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysPro 200
DB 541 ATCAGAGAAATCCCTCAGTACGTACGACAAAGACATCCATATATATTAAGCGCAACCT 600
QY 201 GlnTyrLysValIleSerAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
DB 601 GAATATTAAGGTGATGATGATGATGCGCCAGGATATGTCMAAGCGATGTGGCACAGAC 660
QY 221 SerTyrGlyLeuTyrGlnGlnGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
DB 661 AGTACGCTTGTATGAGCAAGCGCAGATTGTGCGAGTTGCCGATCTGAGTTGGATACA 720
QY 241 GlyArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysIleThrAlaLeuTyrAla 260
DB 721 GGAAGAAACGACGTTTCGATGACATGAAAGCCCTCCGCGTAAATACACGACTATATGCA 780
QY 261 LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
DB 781 CTGGGTGAGCGAATATGCGAATGATACGAACGTCATGATCCCATGTGGCAGGTTCC 840
QY 281 ValLeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
DB 841 GTATTAGAAATGCGCGAAGCAATTAAGAAATGCGACCTCAAGCGAATCTGTTTCCA 900
QY 301 SerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLysPhe 320
DB 901 TCCATCATGATGATGACGTGGTGGCTTGAGGCTTCCTCCATCTGCAAACTTATTC 960
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsn 340
DB 961 AGCCAAAGCATTCAGTCAAGGTGCGCAAAATTCATCAAACTCCGCGGGGCGAGCGTAAAT 1020
QY 341 GlyAlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThr 360
DB 1021 GGGGCTTACACGACATTCACAAATGTGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 361 IleLeuPheAlaAlaGlyAsnGlnArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
DB 1081 ATCTTTTCCGCGCTGGAATGAAGCGCGAAGCGGGGTACATCATGTCAGCTGTGATCG 1140
QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGlnAsnLeuArgProSerPheGlySerTyr 400
DB 1141 GCTAAACAGCCATTAACAGTGGCGCAACCGAAACCTGGTTCAGAGCTTCGTTCTAT 1200
QY 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArg 420
DB 1201 GCAGATATATTAACAGTTCGACAGTTCTTCCCGTGGCCGACAAAGATGGCGCA 1260
QY 421 IleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAla 440
DB 1261 ATCAACCTGATGATGATGCGCGCAGAGGACATACATTTTACACAAAGATCTCTTGGCA 1320
QY 441 ProAspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSer 460
DB 1321 CCGAATTCCTCTTCTGCGGAGATGACAGCAAAATATGCTCATATGGTGGAAAGCTCC 1380
QY 461 MetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValLysAsn 480
DB 1381 ATGGCAACACCATTTTTCGGGAAATGTTTGACAGCTCCGTAGACATTTTGGAAAAAT 1440
QY 481 ArgGlyIleThrProLysProSerLeuLeuLysAlaAlaLeuIleAlaGlyAlaAlaAsp 500
DB 1441 AGAGGAATCACTCTCAAGCTTCCCTATGAAGACAGCTTGAATGCAAGTGTGCTGAT 1500
QY 501 ValGlyLeuGlyTyrArgProAsnGlyAsnGlnGlyTyrGlyArgValThrLeuAspLysSer 520
DB 1501 GTTGAATGGGTATATCGAAGCAAGAAACCAAGATGGGCGCAAGTACCTCGAATTAATCG 1560
QY 521 LeuAsnValAlaTyrValAsnGlnSerSerAlaLeuSerThrSerGlnLysAlaThrTyr 540
DB 1561 TTGAACGTTGCCATATGATGAACGATCCAGTCCCTTCAACTAGCCAAAGAGCAATAT 1620

QY 541 ThrPhetThralaThralaGlyLeuProLeuLeuLeuSerLeuValTTPSerAapAlaPro 560
Db 1621 ACCTTAACTGCAACGGGGGCAACCATTAATAATCCCTGGATAGTGGATGCCCTT 1680
QY 561 AlaSerThrThralaSerValThrlLeuValAsnAspLeuAspLeuValIleThrAlaPro 580
Db 1681 GCAGACATACCTCTCTCTGTAACCTGGTGAATGATTGGATTTGGCTTACAGACCA 1740
QY 581 AenGlyThraGlyrValGlyAsnAspPheSerAlaProPheAspAsnAsnTPaerGly 600
Db 1741 AACGGAACAGATATGTCGGGAATGACTTCTCAGACCATTTGACATATACCTGGATGCG 1800
QY 601 ArgAsnAsnValGlyLeuAsnValPheIleAsnSerProGlnSerGlyThrTyrTrlIleGlu 620
Db 1801 CGGAATTAACGTAAATAATGATTATTAATTCGCCCAAGTGAACATATACCATTTGAG 1860
QY 621 ValGlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1861 GTGCAGCATATATATGTCGGCTTGGACCAATAAATCTCTCGTTGGCAATTTGTAAAC 1917

RESULT 3
US-09-920-954-7
Sequence 7, Application US/09920954
Publication No. US20020064854A1
GENERAL INFORMATION:
APPLICANT: TAKIWA, MIKO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/09/920,954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509,814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-920-954-7

Alignment Scores:
Pred. No.: 7,58e-270 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-7 (1-1923)

QY 2 ArgGlyValGlyValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAGAAAAAGGTGTTTATCTGTTTATCATCGTCGACGATTTTGGCATGTTGGC 66
QY 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLeuGlyIleGln 41
Db 67 TTAAGTAAATCAATCGACAGTGTGTGCAGAGAAATTTTGATCTGATTTCAAGAAATTCAG 126

QY 42 ThrThrThrAspValSerGlyPheSerIyGlnArgGlnThrGlyAlaAlaPheLeu 61
Db 127 ACAACAATGATGCTAAAGGTTTCTTCAAGCAGGGGCAACATGCTGCTCTTTCG 186
QY 62 ValGlnSerGlyAsnValIleLeuLeuLeuGlyLeuLeuLeuLeuValIleThrValPro 81
Db 187 GTGGAATCGAAAATGTGAACCTCCCAAAAGGTTTGCAGAAAGACCTTGAACATGCCCG 246
QY 82 AlaAsnAsnLeuLeuIleValGlnPheAsnGlyProIleLeuGlnGlyThrIyGln 101
Db 247 GCAAATTAATAATCTCATATTAATTCATTAATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGlnThrThrGlyValAlaValIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGAAAAAACAAGGGCAAGATTCGACATACATACCTGATTAATGCTTAATGTC 366
QY 122 GlyTyrGlnGlyAspValGlnSerIyAsnValArgSerIleGlnIleValGlnSerValGlu 141
Db 367 GAGTATGAGGGCGGATGTTTAAGTCAGACACAGACACATTTAGACCTGGATCCGTGAG 426
QY 142 ProTyrLeuProLysTyrIleAspProGlnLeuPheThrLysGlyValaSerThrLeu 161
Db 427 CCTTATTTGCCGATATACGAATATAGTCCCGCTTTTCACAAAGGGGCATCAGAGCTT 486
QY 162 ValLysAlaLeuAlaLeuAspThrIyGlnAsnAsnValGlnIleLeuArgGlyIle 181
Db 487 GTAAAGACGTGGCTTGATACAAAGCAAAATAAAGAGTCAATTAAGAGGCAATC 546
QY 182 GlnGlnIleAlaGlnTyrValaLaserAsnAspValIleTyrIleThrAlaLysProGlu 201
Db 547 GAACAATCGCAATTTGCAATTAAGCAATATGATGCTATATTAACGCAAAAGCTGAG 606
QY 202 TyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSer 221
Db 607 TATAAGGTGATGAATGATGTTGCGGTGAATTCGAAGGAGTGTGCTCAGAGCAGC 666
QY 222 TyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTATGACAAAGCAAGATCGTAGGCTTGCAGATACAGGGCTTGATACAGCT 726
QY 242 ArgAsnAspSerSerMetHisGlyAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeu 261
Db 727 CGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
QY 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyIleGlyThrIleValAlaGlySerVal 281
Db 787 GGACGAGCAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSer 301
Db 847 TTAGAAACGGCTCCACTATTAAGAAATGGCGCTCAGGCGAATCTAGCTTCCAAATCT 906
QY 302 IleMetAspSerSerGlyIleGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSer 321
Db 907 ATCATGATAGTACGGGGGAGCTTGCAGAGCTACCTTGGAATTCGAAACCTTATTCAGC 966
QY 322 GlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsnGly 341
Db 967 CAAGATATCAATGTCTGAGTGCAGAAATTCATCAAACTCTGGGAGCGAGTGAATGGG 1026
QY 342 AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIle 361
Db 1027 GCTTACACACAGATTCAGAAATGTGATATCTATGTCGCCCAAAATGATATGACATTC 1086
QY 362 LeuPheAlaAlaGlyAsnGlyLysArgProAsnGlyGlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGCTGCGGGAATGAAGACCGAAACGGCGAATCATCAGTGCACGAGGACAGCT 1146
QY 382 LysAsnAlaIleThrValGlyAlaThrGlyLeuAsnLeuAspProSerPheGlySerTyrAla 401
Db 1147 AAAAATGCAATTAACGTGCGAGCTTACGGAACCTCCGCCAATCTTGGGTCTTATGCG 1206
QY 402 AspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIle 421

```
Db 1207 GACAAATTCACACATGTCGACAGTTCCTTCACTGACGACCAAGATGACGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTrpIleLeuSerAlaArgSerLeuAlaPro 441
Db 1267 AACCGAGATGTCATGCGACCGGAAACGTTCACTATCAGCAAGATCTTCTTCCGACCG 1326
Qy 442 AspSerSerPheTrpAlaAsnHisAspSerIleValAlaIleValAlaIleValAla 461
Db 1327 GATTCCTCTCTTCTGGGCGAACATGACAGTAAATATGATACATGCGGAGACCTCCATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIleAsnArg 481
Db 1387 GCTACACCGATCGTTGCTGGAACGTGGCAAGCTTCGTGACATTTTGTGAAAAACGA 1446
Qy 482 GlyIleThrProIleProSerLeuLeuValAlaIleValAlaIleValAlaIleValAla 501
Db 1447 GGCATTCACACCAAGACCTTCTCTATTAACCGGACATGATTCGCGGCGCTGACATC 1506
Qy 502 GlyLeuGlyIleProAsnGlyAsnGlnGlyTrpGlyArgValThrLeuAspIleSerLeu 521
Db 1507 GGCCTTGGCTACCCGAAACGATACCAAGATGGGAGACATGACATTTGATTAATCCCTG 1566
Qy 522 AsnValAlaIleValAlaGlnIleSerSerAlaLeuSerThrSerGlnIleValAlaThr 541
Db 1567 AACCTTCCCTATGTAAGACAGATCCAGTTCTCTATCCACAGCAAAAAGGACGATCTCG 1626
Qy 542 PheThrAlaThrAlaGlyIleProLeuIleIleSerLeuValIleSerAspAlaProAla 561
Db 1627 TTAACTGCTACGCGGCGCAAGCTTTGAAAAATCCCTGATAGTGTGATGCCCCCTGCG 1686
Qy 562 SerThrThrAlaSerValThrLeuValAlaAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCACAACCTCTTCGTAAGCTTGTCAATGATGTGACCTTGTCAATACCGCTCCAAAT 1746
Qy 582 GlyThrArgTrpValGlyAsnAspPheSerAlaLeuSerThrPheAspAsnAsnTrpAsp 601
Db 1747 GGCACACAGTATGTAAGAAATGACTTATCTTCGCAATACATGATTAATCGGAGATGCGC 1806
Qy 602 AsnAsnValGlnLeuValPheIleAsnSerProGlnSerGlyThrTrpIleGlnVal 621
Db 1807 AAATACCTAGAAAAATGATTTATTAATGACACCAAGCGGACGTTATACATTTGAAGTA 1866
Qy 622 GlnAspIleValAlaProValGlyProGlnAspPheSerLeuAlaIleValAla 639
Db 1867 CAGGCTTATACGTAACCGGTTGACCAAGAACTTCTGTTGGCAATGTGAAT 1920
RESULT 4
US-10-784-870-7
; Sequence 7, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGETAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/10/784, 870
; PRIOR FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
```

```
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-10-784-870-7
Alignment Scores:
Pred. No.: 7.58e-270 Length: 1923
Score: 3072.00 Matches: 596
Percent Similarity: 97.0% Conservative: 23
Best Local Similarity: 93.4% Mismatches: 19
Query Match: 94.0% Indels: 0
DB: Gaps: 0
US-10-784-870-4 (1-639) x US-10-784-870-7 (1-1923)
Qy 2 ArgIleValIleValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21
Db 7 AAGAAAGAAAAGTGTATATCTGTTATACGCTCAGCGATTTGTGACGTGGCG 66
Qy 22 LeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheIleGln 41
Db 67 TTAAAGTATCATCTGACAGGTGGTCAAGGAATTTGATCTGATTTCAAGGAATTCAG 126
Qy 42 ThrThrThrAspValSerGlyPheSerIleGlnArgGlnThrGlyAlaAlaIlePheLeu 61
Db 127 ACAACAATGATGCTAAAGGTTTCTCCAAACAGGCGACAGCTGCTGCTGCTTTCTG 186
Qy 62 ValGlnSerGlnLeuValIleLeuLeuValGlyLeuLeuValIleValIleValIlePro 81
Db 187 GTGGAATCTAAATGTAAGTAACTCCAAAGGTTTGAGAAAGAGCTTGAACGTCGCG 246
Qy 82 AlaAsnAsnIleLeuHisIleValAlaPheAsnGlyProIleLeuGlnGlnThrIleGln 101
Db 247 GCAATATATATATCATATATATATATATATATATATATATATATATATATATATATAT 306
Qy 102 LysLeuGlnIleThrThrGlyAlaValIleLeuAspTrpIleProAspTrpAlaIleVal 121
Db 307 CAGCTGGAAGAAAACAGGGCGAAAGATTCGACTACATCACTGATTTGCTTACATGTGC 366
Qy 122 GlnIleGlnIleAspValAlaGlnSerIleValArgSerIleGlnIleValIleGlnIle 141
Db 367 GAGATGAGGGCCATGTTAAGTCAAGCAACGACATTAAGACGTCGTCGTCGTCGTCGTC 426
Qy 142 ProIleLeuProIleValIleAspProGlnLeuPheThrIleValIleSerThrLeu 161
Db 427 CTTATTTTGGCGATATATACAAATGATCCCACTTTTCCAAAGGGGCGATCAGAGCTT 486
Qy 162 ValIleAlaLeuAlaLeuAspTrpIleValIleAsnAsnIleValIleGlnLeuArgGly 181
Db 487 GTAAGAGCAGTGGCGCTGATCAAGACGAAATTAAGAGGTGCAATTAAGAGGATC 546
Qy 182 GlnGlnIleAlaGlnIleValAlaIleSerAsnAspValHisIleTrpIleThrAlaVal 201
Db 547 GAACAATGCAACAATGCGCAATTAAGCAATGAGTGTCTATTTTACGCGCAAGCCCTGAG 606
Qy 202 TyrIleValIleMetAsnAspValAlaArgGlyIleValIleValIleAspValAlaGln 221
Db 607 TATTAAGTATGATGATATATATGTCGCTGATATGTCGATGAGGATGTCGTCGTCGTCG 666
Qy 222 TyrGlyLeuTrpIleGlnIleGlnIleValAlaIleValAlaIleAspThrGlyLeuAsp 241
Db 667 TACGGGTGATGACAGGACAGATCGTAGCGGTGCGCATACAGGCGCTTGATACAGGT 726
Qy 242 ArgAsnAspSerSerMetHisIleGlnAlaPheAspGlyIleValIleThrAlaLeuVal 261
Db 727 CGCAATGACAGTTCGATGATGATGAGGCTTCCGCGGGAATTAATCTGATTTATGCAATTG 786
Qy 262 GlyArgThrAsnAsnAlaAsnAspTrpAsnGlyHisGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGACGAATATATCCATATGATACGAATGTCATGTCATGTCATGTCGTCGTCGTCGTA 846
```

```
QY 282 LeuGlyVaNGlyValaThraenlyGlyMeraLaProGlnAlaAenLeuValPheGlnSer 301
Db 847 TTAGAAACGCGCTCACCTAATAAGAAAGCGGCTCAAGGGAATCTAGCTTCCAACT 906
QY 302 IleMetApsSerSerGlyGlyLeuGlyGlyLeuProSerAenLeuGlnThraLeuPheSer 321
Db 907 ATCATGATGACGGGTGGGACTTGGAGGACTACCTTCGAATCTGCACAACTTATTCACG 966
QY 322 GlnAlaPheSerAlaGlyValaArgIleHisThrAenSerTrpGlyValaAlaValaenGly 341
Db 967 CAAGCATACAGTCTGGTGCAGAAATCATCAAACTCTGGGGAGCAGCATGAATGGG 1026
QY 342 AlaThrThrApsSerAraAenValaApsApyrValaArgIysAenApmMetThrIle 361
Db 1027 GCTTACACAAAGATTCAGAAATGTGATGACTATGTGGCAAAATATGATAGCATC 1086
QY 362 LeuPheAlaAlaGlyAenGlyAargProAenGlyGlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTGGCTGCGGGGATGAAAGACCGAACGGCGGAACATCACTGACACAGGACAGCT 1146
QY 382 LysAaHisAlaIleThrValaGlyAlaThrGluAenLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAATCCAAATACAGTCGAGCTACGAAACCTCCGCCCAAGCTTGGGCTTATGCG 1206
QY 402 AspAenIleAenHisValaAlaGlnPheSerSerArgGlyProThrIysAapGlyAargIle 421
Db 1207 GACAAATTCAAACATGTGGCAAGCTTCTTCAAGTACGACGACAAAGAAATGAGCGGATC 1266
QY 422 LysProApyValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGGATGTCAATGGCACCGGGAGCTTCACTACTACACAAAGATCTTCTTGGACCG 1326
QY 442 ApsSerSerThrTrpAlaAenHisApsSerLysTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTCTGGGGGAAACATGACATGAATATGCAATACATGCTGACAGCTCATG 1386
QY 462 AlaThrProIleValaAlaGlyAenValaAlaGluAenArgIleHisPheValaLysAenArg 481
Db 1387 GCTACACCAATCGTGTGGAACGTGGCAAGCTTCTGACAGATTTGTGAAAAACAGA 1446
QY 482 GlyIleThrProLysProSerLeuLeuValaAlaLeuIleAlaGlyValaAlaAapVal 501
Db 1447 GGCATCACACCAAGCTTCTATTTAAAGCGGCACTGATTCGGGTGACGTGACATC 1506
QY 502 GlyLeuGlyTyrProAenGlyAenGlnGlyTyrGlyAargValaThrLeuApyLysSerLeu 521
Db 1507 GGCCTTGGCTAACCGGAACGGTAACAAAGATGGGAGCAGTGAATTTGATTAATCCCTG 1566
QY 522 AenValAlaIleTyrValaAenGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTTGCCATATGTAACGAGTCCAGTTCCTATTCACACGCAAAAGGAGACTGACTCG 1626
QY 542 PheThrAlaThrAlaGlyLysProLeuLysIleSerLeuValIrpSerAapAlaProAla 561
Db 1627 TTTACTGCTACTGCGGCAAGCTTTGAAATCTCCCTGGTATGCTGATGCCCCGCG 1686
QY 562 SerThrThrAlaSerValaThrLeuValaAenApyLeuValaIleThrAlaProAen 581
Db 1687 AGCACAACGCTCTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATCCGCTCCAAAT 1746
QY 582 GlyThrArgTyrValaGlyAenApsPheSerAlaProPheApsAenTrpApsGlyAarg 601
Db 1747 GGCACACAGATATAGAAATGACTTACTTCGCCCATACATATATACCTGGAATGGCCCG 1806
QY 602 AenAenValaGluAenValaPheIleAenSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATTAAGTGAATAATATATTATTAATGACACCAAGCGGAGCGATACATTAATGAAGTA 1866
QY 622 GlnAlaTyrAenValaProValaGlyProGlnAenPheSerLeuAlaIleValaen 639
Db 1867 CAGGCTTAATACGTACCGGTGAGCACAGAACTTCTGTTGGCAATTTGAAAT 1920
```

```
RESULT 5
US-09-920-954-5
; Sequence 5, Application US/09920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUSHISA
; APPLICANT: KOBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-5

Alignment Scores:
Pred. No.: 2,67e-269 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 22
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 3 Gaps: 0

US-10-784-870-4 (1-639) x US-09-920-954-5 (1-1923)

QY 2 ArglyblyblyValaPheLeuSerValaLeuSerAlaAlaAlaIleLeuSerThrValaAla 21
Db 7 AAGAAAGAAAGAGTGTATTTATCTGTATTCAGCTGACGATTTTTCGACTGTGGC 66
QY 22 LeuAenApsProSerAlaGlyApsAlaArgThrPheApyLeuApsPheLysGlyIleGln 41
Db 67 TTAAGTAACTCATCTGCGAGGTGTGCAAGAAATTTGATCTGGAATTTCAAGGAATTCAG 126
QY 42 ThrThrThrApsValaSerGlyPheSerLysGlnArgGlnThrGlyValaAlaAlaPheLeu 61
Db 127 ACAACAACGATGCTAAAGGTTTCTCAAGCAGGGGACAGACTGTGCTGCTTTCTG 186
QY 62 ValGlySerGluAenValaLysLeuLeuLysGlyLeuLeuLysblyblyLeuGluThrValaPro 81
Db 187 GTGAATCTGAAGAAATGTGAACCTCCCAAGGTTTGCAGAAAGAAAGCTTGAACAGTCCG 246
QY 82 AlaAenAenLysLeuHisIleValaGlnPheAenGlyProIleLeuGluGluThrLysGln 101
Db 247 GCAAAATTAATACCTCATATTTTCCAAATTCATAGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuGluThrThrGlyAlaLysIleLeuApyTyrIleProApyTyrAlaTyrIleVala 121
Db 307 CAGCTGAAAGAAAGAGGGGCAAGAAATCTCGACTACATACCTGATTAAGCTTACATTCG 366
QY 122 GlyTyrGlyGlyApsValaGlnSerLysValaArgSerIleGluHisLeuValaGluSerValaGlu 141
Db 367 GAGTATGAGGGGAGATGTTAATGTCAGCAACAGACACATTTGAGCAGTGGAAATCCCGTAG 426
QY 142 ProTyrLeuProLysTyrLysIleApsProGlnLeuPheThrLysGlyAlaSerThrLeu 161
```

D	427	CCTTATTTTCCGATATACAGATTAATCCCGAGCTTTCCAAAGGGGGCATCAGACTT	486
O	162	ValLySaIaLeuAlaLeuAspThrLySGlnAsnAsnLySGluValGlnLeuArgLyIle	181
D	487	GTAAAGCAGTGGCCGCTGATACAAAGCAGAAATTAAGAAGGTGCATTAAAGGCAATC	546
O	182	GIuGIuIIeAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaIysProGlu	201
D	547	GAACAAATTCGCACCAATTGCGCAATMAAGCAATGATATATTAACGGCAAGCCTTAG	606
O	202	TyrLySValMetAsnAspValAlaArgLyIleValLySaIAspValAlaGlnSerSer	221
D	607	TTTAAAGTGCATCATGATGTTGGCGGTGAATTGTCCAAAGCGAATGTGCTTCAGAGCAC	666
O	222	TyrGIyLeuTyrGIyGlnGIyGlnIleValAlaValAlaAspThrGIyLeuAspThrGIy	241
D	667	TACGGGTTGTATGACAAAGACAGATCCGATGGCGTTGCGATACAGGGCTTGATACAGT	726
O	242	ArgAsnAspSerSerMetHisGlnAlaPheArgLyIySIIeThrAlaLeuTyrAlaLeu	261
D	727	CGCATATGACGTTTCATCATGACATGACCCTTCGCGGGAATTAACGCAATTATATGCATTG	786
O	262	GIyArgThrAsnAsnAlaAsnAspThrAsnGIySIIeThrHisValAlaGIIySerVal	281
D	787	GGACCGAAGAAATTAATGCCAATGATTCGAATGTCAATGATACAGATGTGAGCTCCCTA	846
O	282	LeuGIyAsnGIyAlaThrAsnLySGIIyMetAlaProGlnAlaAsnLeuValPheGlnSer	301
D	847	TTAGGAACCGGCTCCACTATTAAGAAATGGGCGCTCAGCGCAATCTAGTCTTCCAACT	906
O	302	IIeMetAspSerSerGIyGIyLeuGIyGlyLeuProSerAsnLeuGlnThrLeuPheSer	321
D	907	ATCATGGATAGCGGTGGGGGAACTTGGAAGAACTACTCGAATCTGCAACCTTATTCAGC	966
O	322	GlnAlaPheSerAlaGIIyAlaArgIleHisThrAsnSerTrpGIIyAlaAlaAsnGIy	341
D	967	CAGCATACAGTGGCTGGTGCACGAATTCATCAAACTCTGGGGAGCAGATGAATAGG	1026
O	342	AlaTyrThrThrAspSerArgAsnValAspAspTyrValArgGIyAsnAspMetThrIle	361
D	1027	GCTTACACAAACGATTCAGAAATGTGATGACTATGTGGCCAAATAATGATATACGATC	1086
O	362	LeuPheAlaAlaGIIyAsnGluArgProAsnGIyGIyThrIleSerAlaProGIIyThrAla	381
D	1087	CTTTTCGCTGCGGGAAAGAAAGAACCGGCGGAACATCAGTGCACAGGCAACACT	1146
O	382	LySaAsnAlaIIeThrValGIyAlaThrGluAsnLeuAspProSerPheGIIySerTyrAla	401
D	1147	AAAAATGCAATPAACGTGGAGCTACGAAAAACCTCCCGCAAGCTTGTGGATCTTATTCG	1206
O	402	AspAsnIIeAsnHisValAlaGlnPheSerSerArgGIyProThrLyAspGIIyArgIle	421
D	1207	GACATATCAACCATGTGGCAACGTTCTTCTTCACGTGACCGAACAAAGATGACGAGATC	1266
O	422	LySProAspValMetAlaProGIIyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro	441
D	1267	AAACCGATGTCATGGCAACGGGAACGTTCAATCATACGAAAGATCTTCTTGGACCG	1326
O	442	AspSerSerPheTrpAlaAsnHisAspSerLySIIyTyrAlaTyrMetGIyGIyThrSerMet	461
D	1327	GATTCCTCTCTCTGGCGCAACCATGACGTAAATATGATCATCATGGGTGGAAACGTCATG	1386
O	462	AlaThrProIIeValAlaGIIyAsnValAlaGlnLeuAspGIIyHisAlaPheValLySaAsnArg	481
D	1387	GCTACACCGAATCGTTGCTGGAAACGTGGCAAGCTTCGTGAGCAATTTGTGAAAAACGA	1446
O	482	GIyIIeThrProLySProSerLeuLeuLySaIaLeuAlaLeuAlaGIIyAlaAlaAspVal	501
D	1447	GGCATCAACCAAAAGCTTCTCTATTAAGAAGGCAATGTCGGGTGGAGCTGACATC	1506
O	502	GIyIleHisGIyTyrProAsnGIyAsnGlnGIyTrpGIyArgValThrLeuAspLySIIySerLeu	521

Db	1507	GGCCTTGCGTACCCGAAACGGTAACCAAGANTGGGAGAGTGAACATTGGANMAATCCCTG	1506
Qy	522	AenValAlaTyrValaAngIuSeSerAlaLeuSerThiSerGlnYsaIaThrTyrThr	541
Db	1567	AACGTGGCTATGTGAACGAGTCCAGTTCTTATCCACGACCAAAAAGCGACCTACTCG	1626
Qy	542	PheThrIaIaThrIaIaGlyIysProLeuLysIleSerLeuValItpSerAapIaIaProIa	561
Db	1627	TTTACTGCTACATGCGCGCAAGCCTTTTAAAAATCCCTGGTAATGCTGANTGCCCCGCG	1686
Qy	562	SerThrThraIaSeSerValThrLeuValaAsnAspLeuAspLeuValIleThrIaIaProAsn	581
Db	1687	AGCAACAAGCTGCTCCGTAACGCTTGTCAAGATGTGAACCTTGCTGATTAACGCTCCAAAT	1746
Qy	582	GlyThrArgTyrValaGlyaAsnAspPheSeSerAlaIaProPheAspAsnItpAspGlyArg	601
Db	1747	GGCACACAGTATGTAGAAATGACTTATTAATCTTCGCATACATATGATTACTGGAGATGCCGC	1806
Qy	602	AsnAsnValaGluAsnValaPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal	621
Db	1807	AATACACGTAGAAATGATTAATTAATTAATGACACCAACAAAGCGGACGTATTAACAATTGAGCTA	1866
Qy	622	GlnAlaIaTyrAsnValaProValaGlyProGlnAsnPheSerLeuAlaIaIaValaAsn	639
Db	1867	CAGGCTTAATACGTACCGGTTGACCAACAAGACTTCTGCTTGCAATTTGAAAT	1920
RESULT 6			
US-10-456-479-3			
Sequence 3, Application US/10456479			
Publication No. US20040072321A1			
GENERAL INFORMATION:			
APPLICANT: SATO, TSUYOSHI			
APPLICANT: OKUDA, MITSUYOSHI			
APPLICANT: TAKIMURA, YASUSHI			
APPLICANT: SUMITOMO, NOBUYUKI			
APPLICANT: NOMURA, MASATAMI			
APPLICANT: KOBAYASHI, TOHRU			
TITLE OF INVENTION: ALKALINE PROTEASE			
FILE REFERENCE: 238700USO			
CURRENT APPLICATION NUMBER: US/10/456, 479			
CURRENT FILING DATE: 2003-06-09			
PRIOR APPLICATION NUMBER: JP 2002-186887			
PRIOR FILING DATE: 2002-06-26			
PRIOR APPLICATION NUMBER: JP 2002-304232			
PRIOR FILING DATE: 2002-10-18			
NUMBER OF SEQ ID NOS: 16			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 3			
LENGTH: 1923			
TYPE: DNA			
ORGANISM: Bacillus sp. KSM-KP43			
FEATURES:			
NAME/KEY: CDS			
LOCATION: (1)..(1920)			
OTHER INFORMATION:			
US-10-456-479-3			
Alignment Scores:			
Pred. No.: 2,67e-269			
Score: 3066.00			
Percent Similarity: 96.9%			
Best Local Similarity: 93.3%			
Query Match: 93.8%			
DB: 7			
Gaps: 0			
US-10-784-870-4 (1-639) x US-10-456-479-3 (1-1923)			
Qy	2	ArgIyselYelYsValPheLeuSeSerValleuSeSerAlaIaIaIeLeuSeThrValaIa	21
Db	7	AAGAAGAAAAAGGCTTTTATCTGTTTATCAAGCTGACGANTTTTGTGACATGTTGCG	66
Qy	22	LeuAsnAsnProSerAlaGlyIaAspAlaArgThrPheAspLeuAspPheIysGlyIleGln	41

Db 67 TTAAGTAATCATCTCAGGTGGTGCAGGAATTTTGTATCTGGATTTCAAGGAATTCAG 126
Qy 42 ThrThrThrAspValSerGlyPheSerIlyGlnArgGlnThrGlyAlaAlaPheLeu 61
Db 127 ACAACAACGATGCTAAAGGTTTCTCCAGCAGGGGCAAGCTGGTCTGCTTTTCG 186
Qy 62 ValGlySerGlyAsnValIlyLeuLeuIlyGlyLeuLeuIlyPheLeuGlnThrValPro 81
Db 187 GTGGATCTGAAATAATGTGAATCTCCAAAAGTTTCCAGAAAGACTTGAACAGTCCG 246
Qy 82 AlaAspAsnIlyLeuHisIleValGlnPheAsnGlyProIleLeuGlnIlyThrIlyGln 101
Db 247 GCATAATATAAATCTCCATATATTCCAATTCAGTACCAATTTTGAAGAAACAAACAG 306
Qy 102 IlyLeuGlnIlyThrThrGlyAlaIlyIleLeuAspTyrIleProAspTyrAlaTyrIleVal 121
Db 307 CAGCTGGAAAAACAGGGGCAAGATTCGACATACCTGATATTCCTTAACATTCG 366
Qy 122 GlyTyrGlnGlyAspValGlnSerIlyValArgSerIleGlnHisValGlySerValGln 141
Db 367 GAGTATGAGGGCGATGTTAAGTCAGCAACAGCAACATTCGACAGTGGAAATCCGCGAG 426
Qy 142 ProTyrLeuProIlySerIlyIleAspProGlnLeuPheThrIlyGlyAlaSerThrLeu 161
Db 427 CCTTATTTGCCATATACAGAAATAGATCCCAAGCTTTTCAAAAAGGGGCATCAGAGCTT 486
Qy 162 ValIlyValAlaLeuAlaLeuAspThrIlyGlnAsnAsnIlyGlyValGlnLeuArgIlyIle 181
Db 487 GTAAAGACGTGGCGCTGTATACAAAGCAAAAAAAGAGGTGCAATTAAGAGGCAATC 546
Qy 182 GlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaIlyProGln 201
Db 547 GAAACAATCGCACAATTCGCAATTAAGCAATGATGCTATATTTACGGCAAAAGCTTGAG 606
Qy 202 TyrIlyValIleAsnAspValAlaArgGlyIleValIlyAlaAspValAlaGlnSerSer 221
Db 607 TATAAGGTGATGAATGATGTTGGCGGTGGAATTTGCAAAAGCGGATGTGGCTCAGACAC 666
Qy 222 TyrGlyLeuTyrGlyGlnIlyGlnIleValAlaAlaAspThrGlyLeuAspThrGly 241
Db 667 TACGGGTGTATGACAAAGACAGATCGTACGGTGGCGATACAGGGCTTGATACAGGT 726
Qy 242 ArgAsnAspSerSerSerMetHisGlnAlaPheArgIlyIleThrAlaLeuTyrAlaLeu 261
Db 727 CCCATGACAGTTCGATGCAAGCTTCCGGGAAAAATTAATCTGCAATTAATGATGCTG 786
Qy 262 GlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisIleGlyThrHisValAlaGlySerVal 281
Db 787 GAGCGGACCAATTAATGCGCAATGATACGAATGCTGATGCGATGTGGCTGGCTCGTA 846
Qy 282 LeuGlyAsnGlyValaThrAsnIlyGlyMetAlaProGlnAlaAsnLeuValaPheGlnSer 301
Db 847 TTAGAAACGGCTCCCAATATAAGAAAGAGCGCTCAGGCAATCTTACTTCCATCT 906
Qy 302 IleMetAspSerSerGlyIlyLeuGlyIlyLeuProSerAsnLeuGlnThrLeuPheSer 321
Db 907 ATCAAGATAGAGCGGTGGGAGCTTGGAGGACTAATCTTCAATCTGAAACCTTATTCAGC 966
Qy 322 GlnAlaPheSerAlaGlyAlaArgIlyIleHisThrAsnSerTyrGlyAlaAlaValaLeuGly 341
Db 967 CAAGCATACAGTGGTGGTCCAGAAATTCATACAAATCTCGGGGAGACACAGATGAGG 1026
Qy 342 AlaTyrThrThrAspSerArgAsnValaAspAspTyrValaArgIlyAsnAspMetThrIle 361
Db 1027 GCTTACACAAAGATTCAGAAATGAGATGACTATGTCGCAAAAATATATATGACGATC 1086
Qy 362 LeuPheAlaAlaGlyAsnGlyIlyArgProAsnGlyIlyThrIleSerAlaProGlyThrAla 381
Db 1087 CTTTTCGCTCGCGGAAATAGAACCGGAAACCAATCATGTCACACAGGACAGCTT 1146
Qy 382 IlyAsnAlaAlaIleThrValaGlyAlaThrGlyAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAATGCAATAACAGTGGAGCTACGGAACCTTCGCGCAAGCTTTGGTCTTATGCG 1206

Qy 402 AsnAsnIleAsnHisValAlaGlnPheSerSerArgIlyProThrIlyAspGlyArgIle 421
Db 1207 GAAATATCTAACCAATGTGGACAGTCTCTTACGTGGACGCAAAAGATGACGAGATC 1266
Qy 422 IlyProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACGGATGTATGACACCGGGAACGTTCAATCACTATACGAAGATCTTCTTGACAGC 1326
Qy 442 AspSerSerPheTyrAlaAsnHisAspSerIlyValaTyrMetGlyIlyThrSerMet 461
Db 1327 GATTCCTCTCTGGGCGAACCATGACAGTAATATGCAATACATGAGTGGAGAACGTCAG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIlyAsnArg 481
Db 1387 GCTACACGATGCTGTGGAAACGTGGACAGCTTCGAGACATTTGTGAANAACGA 1446
Qy 482 GlyIleThrProIlyProSerLeuLeuIlyAlaIleValAlaGlyValaAlaAspVal 501
Db 1447 GGCATCACACCAACCTTCTTATTAAGCGGACATGATTCGCGTGCACCTGACATC 1506
Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValaThrLeuAspIlySerLeu 521
Db 1507 GGCCTTGGCTACCCGACGATACCAAGGATGGGACGAGTGAATGGATTAATCCCTG 1566
Qy 522 AsnValAlaTyrValaAsnGlySerSerAlaLeuSerThrSerGlnIlyValaAlaThrTyrThr 541
Db 1567 AACGTGCTTATGTAAACAGTCTCTTATCCACCAAGCAAAAGCGATGCTG 1626
Qy 542 PheThrAlaThrAlaGlyIlyAspProLeuIlyIleSerLeuValaTyrSerAspAlaProAla 561
Db 1627 TTTATCGTACTGCGGCAAGCTTTGAANAATCTCCGTGATGTGATGCTCCCTGCG 1686
Qy 562 SerThrThrAlaSerValaThrLeuValaAsnAspLeuAspLeuValaIleThrAlaProAsn 581
Db 1687 AGCAACAACGCTCCGTAACGCTGTCATGATCTGCACTTGTGATTAACCTCCAAT 1746
Qy 582 GlyThrArgTyrValaGlyAsnAspPheSerAlaProPheAspAsnThrAspGlyArg 601
Db 1747 GGCACAGTATGTAGAAATATCTTACTTCCATCAATCAAGTAACTCGGATGGCCGCT 1806
Qy 602 AsnAsnValaGlyAsnValaPheIleAsnSerProGlnSerGlyTyrThrIleGlyVala 621
Db 1807 AATAACGTAGAAATATTTATTAATGACACCAAGGGGACGTATACATTAAGCTGA 1866
Qy 622 GlnAlaTyrAsnValaProValaGlyProGlnAsnPheSerLeuAlaIleValaAsn 639
Db 1867 CAGGCTTATAACGTACCGGTGGACCAAGACCTTCTCGTGGCAATGTGAAT 1920
RESULT 7
US-10-784-870-5
Sequence 5, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIRATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784, 870
PRIOR APPLICATION NUMBER: 2004-02-24
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24

QY 502 G1YLeuG1YrProaEng1YAsnG1YrProG1YrValThrLeuAspIleSerLeu 521
DB 1507 GGCTTGGCTACCGGAAACGTAACGATGGGAGAGATGACATTAATGATCCCTG 1566
QY 522 AsnValAlaThrValaAngIleuSerSerAlaLeuSerThrSerGlnYValaThrThr 541
DB 1567 AACGTCCTATGATGAGAGATCCAGTCTCTATCCACGACCAAAAAGAGATCTCG 1626
QY 542 PheThrAlaThrAlaGlyLeuProLeuYValSerLeuValTrpSerAspAlaProAla 561
DB 1627 TTTACTGCTACTGCGGCAACCTTTGAAATCTCCCTGATGCTGATGCCCCCTGCG 1686
QY 562 SerThrAlaSerValThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
DB 1687 AGCAACCTGCTCCGTAACGCTTGCAATGATGAGACCTTGTCATTAACGCTCAAT 1746
QY 582 G1YrArgTrpValG1YAsnAspPheSerAlaProPheAspAsnAntTrpAspG1YArg 601
DB 1747 GGCACACAGATGAGAAATGACTTACTTCGCCATACATGATTAATGGGATGGCCG 1806
QY 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerG1YrThrTrpIleGluVal 621
DB 1807 AATAACCTAGAAATGATTAATTAATGACACCAAAAGCGGAGATTAATTAATGAGTA 1866
QY 622 GluAlaThrAsnValProValG1YrProGlnAsnPheSerLeuAlaIleValAsn 639
DB 1867 CAGGCTTAATACGTAACGCTTGACACACACACACCTTCTGTTGGCAATTTGAAAT 1920
RESULT 9
US-10-820-714A-2
Sequence 2, Application US/10820714A
Publication No. US20050214922A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Kobayashi, Tohru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasuhiro
APPLICANT: Sato, Teiyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697USO
CURRENT APPLICATION NUMBER: US/10/820,714A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: (1)..(618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619)..()
US-10-820-714A-2
Alignment Scores:
Pred. No.: 2,67e-269 Length: 1923
Score: 3066.00 Matches: 595
Percent Similarity: 96.9% Conservative: 23
Best Local Similarity: 93.3% Mismatches: 20
Query Match: 93.8% Indels: 0
DB: 9 Gaps: 0
US-10-784-870-4 (1-639) x US-10-820-714A-2 (1-1923)
QY 2 ArgLysLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrValAla 21

DB 7 AAGAGAAAAGAGGTTTATCTGTTTATCTGCTGACGACGATTTGTGACTGTCG 66
QY 22 LeuAsnAspProSerAlaG1YAspAlaArgThrPheAspLeuAspPheYsg1YleGln 41
DB 67 TTAAGTAATCATCTGAGGTGGGCAAGGAATTTGATCTGGATTTCAAGGAATTCAG 126
QY 42 ThrThrThrAspValSerG1YPheSerLysGlnArgGlnThrG1YAlaAlaIlePheLeu 61
DB 127 ACAACAACCTGATCTAAAGGTTTCTCCAAAGCAGGGGAGACTGCTGCTGCTTCTG 186
QY 62 ValG1YSerG1YAsnValLysLeuLeuLysG1YLeuLeuYsgLysLeuG1YrPro 81
DB 187 GTGAAATCTGAAATGTAAGTGAACCTCCAAAGGTTTGCAGAGAGAGCTTGAACAGTCCG 246
QY 82 AlaAsnAsnLysLeuValIleValGlnPheAsnG1YProIleLeuGlnG1YrProLysGln 101
DB 247 GCAAAATATACTCCATATTAATCAATTCATCAATGACCAATTTTGAAGAAACAAACAG 306
QY 102 LysLeuG1YrThrThrG1YAlaLysIleLeuAspTrpIleProAspTrpAlaTrpIleVal 121
DB 307 CAGCTGGAAGAAACAGGGGCAAGATTCGACTACATACCTGATTAATGCTTACATTTGC 366
QY 122 GluTrpGluG1YAspValGlnSerLysValArgSerIleGluIleValG1YSerValGlu 141
DB 367 GAGTATGAGGGCGATGTAAAGTCAACCAAGCAACCAATTAAGACCTGGAATCCGTGGAG 426
QY 142 ProTrpLeuProLysTrpLysIleAspProGlnLeuPheThrLysG1YAlaSerThrLeu 161
DB 427 CTTATTTGCCGATATACAGAAATGATCCCACTTTTCAAAAAGGGGATCGAGCTT 486
QY 162 ValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGluValGlnLeuArgG1Yle 181
DB 487 GTAAAGACAGTGGCGCTTGATCAAGACAGAAATAAAGAGGTGCAATTAAGAGCATC 546
QY 182 GluGluIleAlaGlnTrpValAlaSerAsnAspValHisTrpIleThrAlaLysProGlu 201
DB 547 GAACAAATCGCAAAATTCGCATTAACCAATGAATGTCTATATTAATCGGCAAGCCCTGAG 606
QY 202 TrpLysValIleAsnAspValAlaArgG1YIleValLysAlaAspValAlaGlnSerSer 221
DB 607 TATTAAGTATGATGATATGATGTCGCTGCAATTTGTCAAAAGCGATGTGCTCGAGACGC 666
QY 222 TrpG1YLeuTrpG1YGlnGlnGlnIleValAlaValAlaAspThrG1YLeuAspThrG1Y 241
DB 667 TACGGGTGTATGACACAGACAGATCGTAGCGGTGCGCATTCAGAGGCTTGATACAGGT 726
QY 242 ArgAsnAspSerSerMetHisGluAlaPheArgG1YLysIleThrAlaLeuTrpAlaLeu 261
DB 727 CGCAATGACAGTTCCGATGATGATGAAGCTTCCCGGGGAAATTAATTCATTAATGCAATG 786
QY 262 G1YArgThrAsnAsnAlaAsnAspThrAsnG1YHisG1YrHisValAlaG1YSerVal 281
DB 787 GAGCGACCAATTAATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATG 846
QY 282 LeuG1YAsnG1YAlaThrAsnLysG1YMetAlaProGlnAlaAsnLeuValPheGlnSer 301
DB 847 TTAGAAGACGGCTCCACTAATAAGGAATGGCGCTCAAGCGCAATCTAGCTTCCAAATCT 966
QY 302 IleMetAspSerSerG1YG1YLeuG1YG1YLeuProSerAsnLeuGlnThrLeuPheSer 321
DB 907 ATCATGATGATGAGGTGGGAGCTTGAGAGACTTCCAAATCTGCAAACTTAATTCACAC 966
QY 322 GlnAlaPheSerAlaG1YAlaArgIleHisTrpAsnSerTrpG1YAlaAlaValaAng1Y 341
DB 967 CAGACATACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
QY 342 AlaTrpThrThrAspSerArgAsnValaAspAspTrpValaArgLysAsnAspMetThrIle 361
DB 1027 GCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
QY 362 LeuPheAlaAlaG1YAsnG1YrProaEng1YG1YrHisSerAlaProG1YrThrAla 381

```
Db 1087 CTTTCGCTGCGGGAATGAAGACCGAGCGGGAACATCATGTGACCGACAGCACT 1146
Qy 382 LysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAla 401
Db 1147 AAAAAAGCAATTAACAGTCGAGCTACGGAAACCTCCGCCCAAGCTTTGGGCTTATGCG 1206
Qy 402 AspAsnIleAsnIleValAlaGlnPheSerArgGlyProThrIysAspGlyArgIle 421
Db 1207 GACAAATATCAACATGTGGGACAGTTCTTTCACGTGGACCGCAACAAAGATGACCGATC 1266
Qy 422 LysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaPro 441
Db 1267 AAACCGGATGTCAATGGACCGGGAACGTTTATCATATCAAGCAAGATCTTCTTGACACG 1326
Qy 442 ArgSerSerPheThrAlaAsnIleAspSerIysTyrAlaTyrMetGlyGlyThrSerMet 461
Db 1327 GATTCCTCTTCGCGGGAACCATGACAGTAATATGCAATACATGGGTGGAAAGTCACATG 1386
Qy 462 AlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluIlePheValLysAsnArg 481
Db 1387 GCTACACCGATCTGCTTGGAAACGTGGACAGCTTCGTGAGCATTTTGTGAAAAACAGA 1446
Qy 482 GlyIleThrProIysProSerIleuLysAlaIleuIleAlaGlyAlaIleAspVal 501
Db 1447 GGCATCACACCAAGGCTTCTCATTTAAAGCGGCACTGATTCGCGGTGACGCTGACATC 1506
Qy 502 GlyLeuGlyTyrProAsnGlyAsnGlnGlyTyrGlyArgValThrIleuAspIysSerLeu 521
Db 1507 GGCCTTGGCTTACCCGACCGATACCAAGAGTGGGACGATGACATGGATTAATCCCTG 1566
Qy 522 AsnValAlaIleTyrValAsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTyrThr 541
Db 1567 AACGTGGCTTAATGTGAACAGATCCAGTCTCTATCCACGACGCAAAAACGACGATCTG 1626
Qy 542 PheThrAlaIleThrAlaGlyIysProLeuLysIleSerLeuValITrSerAspAlaProAla 561
Db 1627 TTACTGCTACTGCGCGGCAAGCTTTGAAAAATCTCCCTGTATGGTCTGATCCCTCGG 1686
Qy 562 SerThrThrAlaSerValThrIleuValAsnAspLeuAspLeuValIleThrAlaProAsn 581
Db 1687 AGCACAACGCTTCCGTACCGCTTGCATGATCTGGACCTTGCTTACCCCTCCAAAT 1746
Qy 582 GlyThrArgTyrValGlyAsnAspPheSerAlaProPheAspAsnITrAspGlyArg 601
Db 1747 GGCACACAGTATGTAGGAATGATCTTACTTCGCCATACATGATGATGGATGGCCG 1806
Qy 602 AsnAsnValGluAsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGluVal 621
Db 1807 AATTAACGTAGAAATGTATTATTAATGACACCAACAAAGCGGACGATATCAATTTGAGTA 1866
Qy 622 GlnAlaTyrAsnValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
Db 1867 CAGGCTTATTAACGTACCGGTTGGACCAAGACCTTCTGTTGGCAATTTGTAAAT 1920
RESULT 10
US-10-820-712A-11
; Sequence 11, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Izawa, Yoshihumi
; APPLICANT: Kobayashi, Toru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tatsuoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: P51701-USO
; CURRENT FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; PRIOR FILING DATE: 2003-04-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 11
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-KP9860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: protease KP9860
US-10-820-712A-11
Alignment Scores:
Pred. No.: 6,696-194 Length: 1302
Score: 2237.00 Matches: 433
Percent Similarity: 99.88 Conservative: 0
Best Local Similarity: 99.88 Mismatches: 1
Query Match: 68.54 Indels: 0
DB: 8 Gaps: 0
US-10-784-870-4 (1-639) x US-10-820-712A-11 (1-1302)
Qy 206 AsnAspValAlaArgGlyIleValIysAlaAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db 1 AATGATGTGGCCAGAGTATGTGAAACCGGATGTGGACAGACGCTACGGTTGTAT 60
Qy 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAspAspSer 245
Db 61 GGCACAGGCGAGATGTGCGAGTTGCCATCTGGATTGGATGACGAAAGAACGACAGT 120
Qy 246 SerMetIleGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db 121 TCGATGCATGAAGCTTCGCGGTAAATTAACAGCATATATGACCTGGGTGCGACGAT 180
Qy 266 AsnAlaAsnAspThrAsnGlyIleGlyThrHisValAlaGlySerValLeuGlyAsnGly 285
Db 181 AATGGAATGATACGAAGGTCATGTGATCCCATGTGGCAGGTTGGTATTAAGAAATGCG 240
Qy 286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
Db 241 GCACGAAATTAAGAAATGACCTTCAGCCAAATCTGGTTTTCATCATCATCATGATATGC 300
Qy 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrIleuPheSerGlnAlaPheSer 325
Db 301 AGTGGTGGCTTGGAGGCTTCCATTCATTCGCAACCTTATTCAACGCAAGATTGAGT 360
Qy 326 AlaGlyAlaArgIleHisThrAsnSerITrGlyAlaAlaValAsnGlyAlaTyrThrThr 345
Db 361 GCAGGTGCGAATTCATTAACAACTCTGGGGGCGACGCGGATGGGGCTTACACAGCA 420
Qy 346 AspSerArgAsnValAspAspTyrValArgLysAsnAspMetThrIleuPheAlaIle 365
Db 421 GATTCGAAATGTGATGATGATCTATTAAGAAATATATATGACATTTCTTTGCGGCT 480
Qy 366 GlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLysAsnAlaIle 385
Db 481 GGGAAATGAAGGCGGAACGCGGCTACCATCATGCTCACTGAGCTTAAAAAAGCCATC 540
Qy 386 ThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
Db 541 ACTGTGGCGCAACCGAAACCTGGCTCAAGCTTCGGTTCTTATGAGATTAATTAAC 600
Qy 406 HisValAlaGlnPheSerSerArgGlyProThrIysAspGlyArgIleLysProAspVal 425
Db 601 CAGCTTGACAGTTCTTCTCCCGTGGGCCGACAAABAATGGCGAATCAAGCTGATGTC 660
Qy 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
Db 661 ATGGCGCCAGGCAATCATTTATCAAGAAAGCTTCTTCGACCCGATTCCTCTTC 720
Qy 446 TrpAlaAsnHisAspSerIysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
Db 721 TGGCGAATCATGACCAAAATATGCTTATATGGGTGAACGTCATGGCAACCGCAAT 780
Qy 466 ValAlaGlyAsnValAlaGlnLeuArgGluIlePheValLysAsnArgGlyIleThrPro 485
```

Db 781 GTGGGGGAAATGTGCACAGCTCCGTGAGCATTTTGAAAAATAGAGGAAATCACTCCT 840
Qy
Qy 486 LysProSerLeuLeuLysAlaAlaLeuLleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db 841 AAGCTTCCCTATTTGAAAGCAGCTTTGATTCAGAGTCTGCTGATGTTGGATTGGCTTAT 900
Qy 506 ProAsnGlyAsnGlnGlyTPGlyArgValThreAsnAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACGAAACCAAGATGGGCGCAGATGACCTGGATTAATCGTTGAACGTTGCCCTAT 960
Qy 526 ValAsnGlySerSerAlaLeuSerThrSerglnLysAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAATCCAGTGCCTTATCACTAGCCAAAACGACATTACTTACTTACGCAAG 1020
Qy 546 AlaGlyLysProLeuLysLleSerLeuValITPserAspAlaProAlaSerThrAla 565
Db 1021 GCGGGCAAGCCATTGAAAATCTCCCTGTATGTCGATGCCCTCGAAGCACTACTGCT 1080
Qy 566 SerValThreValAsnAspLeuAspLeuValIThrrAlaProAsnGlyThrArgTyr 585
Db 1081 TCTGTAACTCTGTGATGATTTGGATTGGTCACTACAGCAACCAAGCAAGTAT 1140
Qy 586 ValGlyAsnAspPheSerAlaProPheAspAsnAsnITrpAspGlyArgAsnAsnValGlu 605
Db 1141 GTGGGGAATGACTTCTCAGACCATTTGACATTAACGGATGGCCCAATTAAGTAA 1200
Qy 606 AsnValPheIleAsnSerProGlnSerglyThrTyrIThrrIleGluValGlnAlaTyrAsn 625
Db 1201 AAGTATTTATTAATTCGCCCAAGTGAACATATACATGAGTGGCAAGCATATAT 1260
Qy 626 ValProValGlyProGlnAsnAspSerLeuAlaIleValAsn 639
Db 1261 GTCCCGGTGGACCAAAATCTCTGTTGCAATGTGAAC 1302

RESULT 11
US-10-820-714A-12
Sequence 12, Application US/10820714A
Publication No. US20050214922A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Kobayashi, Tohru
APPLICANT: Sumitomo, Nobuyuki
APPLICANT: Takimura, Yasuhiro
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251697US0
CURRENT APPLICATION NUMBER: US/10/820,714A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106709
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.2
SEQ ID NO 12
LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus sp. KSM-kp9860
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1302)
OTHER INFORMATION: protease KP9860
US-10-820-714A-12

Alignment Scores:

Pred. No.: 6,69e-194 Length: 1302
Score: 2237.00 Matches: 433
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 68.5% Indels: 0
DB: 9 Gaps: 0

US-10-784-870-4 (1-639) x US-10-820-714A-12 (1-1302)

Qy 206 AsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSerSerTyrGlyLeuTyr 225
Db 1 AATGATGTGGCCAGAGGTATTTGAAAGCCGAATGTGCACAGAGCACTACGGTTTAT 60
Qy 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
Db 61 GGAACAGGCCAGATTGTCCGAGTTGCCGATCTCGAATTTGATACAGGAAGAAACGACGT 120
Qy 246 SerMetHisGlyIleAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
Db 121 TCCATCATGAAACCTTCCCGGTTAAATPAACAGCACTATATCACTGCGGTCCGACGAAT 180
Qy 266 AsnAlaAsnAspThrAsnGlyHisGlyIThrHisValAlaGlySerValLeuGlyAsnGly 285
Db 181 AATGCCAAATGATCGAAGCGGTCAATGATCCCAATGGCAGGTTCCGATTATGAAATAGGC 240
Qy 286 AlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
Db 241 GCACCAATTAAGGAATGGCACTTCAAGCAATCTGGTTTTCATCTCATCATGATAGC 300
Qy 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnIThrLeuPheSerglnAlaPheSer 325
Db 301 AGTGTGGGCTTGAGAGCTTGCTTCCATCTGCAAACTTATTCAGCCAAAGCATTCAGT 360
Qy 326 AlaGlyValaArgIleHisIThrAsnSerITPGLyAlaAlaValaAsnGlyAlaTyrThr 345
Db 361 GCAGGTGCCGAATTCATACAACTCTGGGGGGCAGCGGTGAATGGGGCTTCACACACA 420
Qy 346 AspSerArgAsnValAlaAspPyrValArgLysAsnAspMetThrIleLeuPheAlaAla 365
Db 421 GATTCCGAATATGATGATGATCTATGAGAAATGATATGACGATTTCTTTTCGGGCT 480
Qy 366 GlyAsnGlyLysProAsnGlyGlyIThrIleSerAlaProGlyIThrAlaLysAsnAlaIle 385
Db 481 GGGAAATGAAGGCGCAAGCGCGTACATCATGACAGTGTGACGGCTTAAAGCGCATC 540
Qy 386 ThrValGlyValaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
Db 541 ACTGTGGGCGCAACCAAACTCGCTCCAAAGCTTCGTTCATGCAATATATTAATAC 600
Qy 406 HisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLysProAspVal 425
Db 601 CACGTTGCACAGTCTCTTCCGTGGCCGACAAAGATGGCGAATCAAGCTGATGTC 660
Qy 426 MetAlaProGlyIThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
Db 661 ATGGCCCGAAGGACATACATTTTATCAGCAAGATCTTCTTCCGACCCGATTCCTCCTTC 720
Qy 446 TrpAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyIThrSerMetAlaITrProlle 465
Db 721 TGGGCAATCATACAGCAAAATATGCTTATATGGTGAACGTCCATGCGCAACCGAAT 780
Qy 466 ValAlaGlyAsnValAlaGlnLeuArgGlyHisPheValLysAsnArgGlyIleIThrPro 485
Db 781 GTTGGGGGAATGTGCACAGCTCCGTGAGCATTTGTGAAAATAGAGAAATCACTCCT 840
Qy 486 LysProSerLeuLeuLysAlaAlaLeuLleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
Db 841 AAGCTTCCCTATTTGAAAGCAGCTTTGATTGACAGGTGCTGATTTGGATTGGGTTAT 900
Qy 506 ProAsnGlyAsnGlnGlyITPGLyArgValThreAsnAspLysSerLeuAsnValAlaTyr 525
Db 901 CCGAACGAAACCAAGATGGGCGCAGATGACCTGGATTAATCGTTGAACGTTGCCCTAT 960
Qy 526 ValAsnGlySerSerAlaLeuSerThrSerglnLysAlaThrTyrThrPheThrAlaThr 545
Db 961 GTGAACGAATCCAGTGCCTTATCACTAGCCAAAACGACATTACTTACTTACGCAAG 1020
Qy 546 AlaGlyLysProLeuLysLleSerLeuValITPserAspAlaProAlaSerThrAla 565
Db 1021 GCGGGCAAGCCATTGAAAATCTCCCTGTATGTCGATGCCCTCGAAGCACTACTGCT 1080


```
/ Publication No. US20050214922A1
/ GENERAL INFORMATION:
/ APPLICANT: KAO CORPORATION
/ APPLICANT: Okuda, Mitsuyoshi
/ APPLICANT: Kobayashi, Tohru
/ APPLICANT: Sumitomo, Nobuyuki
/ APPLICANT: Takimura, Yasuhide
/ APPLICANT: Sato, Tsuyoshi
/ TITLE OF INVENTION: ALKALINE PROTEASE
/ FILE REFERENCE: 251697USO
/ CURRENT APPLICATION NUMBER: US/10/820, 714A
/ PRIOR FILING DATE: 2004-04-09
/ PRIOR APPLICATION NUMBER: 2003-106709
/ PRIOR FILING DATE: 2003-04-10
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 14
/ LENGTH: 1302
/ TYPE: DNA
/ ORGANISM: Bacillus sp. KSM-9865
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1302)
/ OTHER INFORMATION: Procease 9865
US-10-820-714A-14

Alignment Scores:
Pred. No.: 1,98e-189 Length: 1302
Score: 2188.00 Matches: 419
Percent Similarity: 99.3% Conservative: 12
Best Local Similarity: 96.5% Mismatches: 3
Query Match: 67.0% Indels: 0
DB: Gaps: 0

US-10-784-870-4 (1-639) x US-10-820-714A-14 (1-1302)

QY 206 AenAspVa1AaArgIy1LeVal1AaSpVa1AaGInSerSerTyrgIyLeuTy 225
DB 1 AATGATGTTTCACTGATGAAATTTGCAAGCGAGTGTGGCGGACGACTTCGGGTTGTAT 60
QY 226 GlyGInGlyGIn1LeVal1Aa1AaSpThrgIyLeuAspThrgIyArGAsnAspSer 245
DB 61 GGAACAAGACAGATCGAGCGGTTGCCGATACAGGGCTTGATACAGGTCCAAATGACAGT 120
QY 246 SerMetH1sgIuA1AaPheArgIyLygIleThrAlaLeuTyrgAlaLeuGlyArGThra 265
DB 121 TCGATGCAATGAAAGCTTCCGGGGAAATTAATGCAATTAATGCAATGGACGACGAAAT 180
QY 266 AenAlAaSnAspThraSngIyH1sgIyThraH1AaGlySerVal1LeuGlyAaGngIy 285
DB 181 AATGCCAAATGATTCGAATGATTCATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 286 AlaThraSngIyMetAlaProGInAlAaSnLeuVal1PheGInSerI1LeuEaPser 305
DB 241 TCCACTAAATAAAGAAATGGGCTCCAGGCGAAATCTAGTCTCCAAATCTAATGATAGC 300
QY 306 SerGlyGlyLeuGlyGlyLeuP1SerAsnLeuGInThraLeuPheSerGInAlaPheSer 325
DB 301 GGTGGGGGACTTGGAGACTTACCTTGAATCTGCAACCTTATTCAGCCAAAGCATACAGT 360
QY 326 AlaGlyAlaArgI1LeH1SerThraSerSerTPGlyAlaAlaVal1AaGngIyAlaTyrrTr 345
DB 361 GCTGGAGCCGAATTCATACAACTCTGAGGAGACGAGCATGATGGGCTTACACAAAC 420
QY 346 AspSerArgAsnVal1AaPspTyrgVal1ArgLyAsnAspMetThrI1LeuPheAla1Aa 365
DB 421 GATTCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATGATCTTTTCCGTGCGC 480
QY 366 GlyH1sgIyArGProaSngIyGlyThrI1SerAlaProGlyThraAlaYaaAla1Aa 385
DB 481 GGGAAATGAAGACCGAAGCGGAAACCATCATGAGCAACGAGGCAAGCTAAATGAATAA 540
QY 386 ThrValGlyAlaThrgIuSngIyAaArgProSerPheGlySerTyrgAlaAspAsnI1Lea 405
```

```
DB 541 ACAAGTCGAGACTACGAAAACTCCGCCCAAGCTTGGGCTTATGCGGACAAATATAC 600
QY 406 H1eVal1AaGInPheSerSerArgIyProThraSngIyArG1yArG1yLeuP1Aa 425
DB 601 CATGTGACAGATTTCTTCTTCACTGAGCCGACAAAGATGAGCGATCAACCGGATGTC 660
QY 426 MetAlaProGlyThrTyrrI1LeuSerAlaArgSerSerI1LeuAlaProAspSerPhe 445
DB 661 ATGGCAACCGGAAAGCTTCATATCATGCAAGATCTTCTTCAACCGGATTCCTCTTC 720
QY 446 TrpAlAaSnH1AaPserTyrgVal1AaTyrgMetGlyGlyThraSerMetAlaThraProI1e 465
DB 721 TGGGCGAACCATACAGTAAATATGATACATGAGTGGAGAGCTTCATGCTTACCGCATC 780
QY 466 Val1AaGlyAaSnVal1AaGInLeuArgIyH1AaPheVal1AaSngIyArG1yLeuPro 485
DB 781 GTTGCTGAAACGTGGACAGCTTCTGAGCATTTTGTGAAAAACAGAGGCATCACACCA 840
QY 486 TygProSerI1LeuLeuTyrgAlaAlaLeuI1AaGlyAlaAlaAspVal1GlyLeuGlyTy 505
DB 841 AAGCTTCTCTATTAAGGCGCATGATTTGCCGAGCAGACATCGGCTTGGCTAC 900
QY 506 ProAsnGlyAaSngIyTrpGlyArGVal1ThraSngIySerI1LeuAsnVal1AaTy 525
DB 901 CCGAACCGTACCAAGATGGGACAGATGACATTTGATTAATCCTTAACGTTGCTTAT 960
QY 526 Val1AaSngIySerAlaLeuSerThraSngIyAlaThrTyrrThraPheThraAlaTh 545
DB 961 GTGAACAGATCGAGTCTTATCCACAGCAAAAGCGACGATCGTTTACGTGACT 1020
QY 546 AlaGlyLygProLeuLygI1SerI1LeuVal1TrpSerAspAlaProAlaSerThraAla 565
DB 1021 GCCGGACGCTTGAAGATCTCCGTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 566 SerVal1ThraLeuVal1AaSnAspLeuAspLeuVal1I1eThraAlaProaSngIyThraG 585
DB 1081 TCGTAAACGCTTCTCATATACCTGAGCTTGTGATTAACGCTTCAATGACACAAATAT 1140
QY 586 Val1AaSngIyPheSerAlaProPheAsnAsnTrpAspGlyArGAsnAsnVal1Gly 605
DB 1141 GTTGAATATGACTTTTACTTCGCAATCAATTAATTAATGAGGATGGCGCAATTAACGTA 1200
QY 606 AenVal1PheI1AaSnSerProGInSerGlyThrTyrrI1GlyVal1GInAlaTyra 625
DB 1201 AATGATTTATTAATTCGCGCAAAAGCGGACGATATGATGATGATGATGATGATGAT 1260
QY 626 ValProVal1GlyProGInAsnPheSerI1LeuAlaI1Val1Aa 639
DB 1261 GTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGTAAC 1302

RESULT 14
US-10-385-662-1
/ Sequence 1, Application US/10385662
/ Publication No. US20040002432A1
/ GENERAL INFORMATION:
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SATO, TSUYOSHI
/ APPLICANT: SATO, KAZUHIRO
/ APPLICANT: SUMITOMO, NOBUYUKI
/ APPLICANT: IZAWA, YOSHIKUNI
/ APPLICANT: SAKETI, KATSUHIISA
/ APPLICANT: KOBAYASHI, TOHRU
/ TITLE OF INVENTION: ALKALINE protease
/ FILE REFERENCE: 234938USO
/ CURRENT APPLICATION NUMBER: US/10/385, 662
/ PRIOR FILING DATE: 2003-03-12
/ PRIOR APPLICATION NUMBER: JP 2002-091428
/ PRIOR FILING DATE: 2002-03-22
/ PRIOR APPLICATION NUMBER: JP 2002-165987
/ PRIOR FILING DATE: 2002-06-06
/ PRIOR APPLICATION NUMBER: JP 2002-304230
```


PRIOR FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: JP 2002-304231
PRIOR FILING DATE: 2002-10-18
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1305
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1305)
OTHER INFORMATION:
US-10-385-662-1

Alignment Scores:

Pred. No.:	5,67e-189	Length:	1305
Score:	2183.00	Matches:	418
Percent Similarity:	99.3%	Conservative:	13
Best Local Similarity:	96.3%	Mismatches:	3
Query Match:	66.8%	Indels:	0
DB:	6	Gaps:	0

US-10-784-870-4 (1-639) x US-10-385-662-1 (1-1305)

QY 206 AaasapvalaiaarglyilevallysaiaasapvalaiaaInserSertYrGlyleuYr 225
DB 1 AATGATGTTGGCGGTGGAATTTGCAGAGCGAGTGGCTCAGAGCAAGTCCGGTTGTAT 60
QY 226 GLyGInGlyGInIleValaIaValaIaAapThrGlyLeuAapThrGlyAraAapSer 245
DB 61 GGAACAAGACAGATCTTACCGGTGCGGATACAGGCGCTTGAATACAGTGCATGACAT 120
QY 246 SerMetHISGLuaIaAphAapRgIlyvSIleThraIaLeuYrIaIeueuIyAraThraa 265
DB 121 TCGATGCAAGAACCTTCGCGGAAATTACTGCAATTAATGATGATGGAGCGAAGAT 180
QY 266 AaanaIaasnaAapThraaNgIyHISGlyThraHISValaIaGlySerValIeueuIyAaNgly 285
DB 181 AATGCCAATGATACGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 286 AIAthrAaenIyGlyMeCaIaProGInaIaAaenIeuaIaPhaGInSerIleMeCaAapSer 305
DB 241 TCCACTTAATTAAGAAATGCGCTCAGGGAATCTAGCTTCAATCTATCATGATGAC 300
QY 306 SerGlyGlyLeuGlyGlyLeuPProSerAaenIeueuIyAaIaPhaSerGInaIaAaPheSer 325
DB 301 GGTGGGGGACTTGGAGGACTTCTTGCATCTGCAAACTTATTCAGCCAGCATACAT 360
QY 326 AIAgIyAIAaGlyIeHISThrAaenSerIIPGlyAIAIAValaIaAaNglyAIAIyThrThra 345
DB 361 GCTGGTGCAGAAATTCATCAAACTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 346 AaPserAraAaenValaAaPserIyYrValaIaAaGlyAaAaPserMetThraIeueuPheAaIa 365
DB 421 GATTCACAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 366 GLyAaNgIyAaGlyProAaNgIyGlyThraIeSerAaIaProGlyThraIaIyAaAaIaIe 385
DB 481 GGGATGAAGACCGAAGCGGAAACCATCATGATGATGATGATGATGATGATGATGATGAT 540
QY 386 ThrValIaGlyAIAthrGlyAaenIeueuAaPProSerPheGlySerYrAIAaPserAaIaAa 405
DB 541 ACAAGTCGAGCTACGAGAAACCTCCGCCCAAGCTTTGGGTCTTATGCGGACAAATATCAC 600
QY 406 HISValaIaGInPheSerSerAraGlyProThraIyAaPserGlyAaGlyIleYbProAaPVal 425
DB 601 CATGTGGACAGTTCTTCTTCAAGTGAACCGAACAAGATGAGCGGATCAAAACCGGATGTC 660
QY 426 MeCaIaPProGlyThraIyIeueuSerAIAaPserSerIeueuAIAProAaPserPhe 445
DB 661 ATGGACCGGAGACTTCACTACTATCAGCAAGATCTTCTTTCACCGGATTCCTCTTC 720

QY 446 TTPaIaAaenHISaPserIyYrAIAIyThraMetGlyGlyThraSerMeCaIaAthrProIle 465
DB 721 TGGGGAACATGACGATGAATATGATATGATGATGATGATGATGATGATGATGATGAT 780
QY 466 ValaIaGlyAaenValaIaGInIeuaRgIuHISPheValIyAaAaRgGlyIleThraPro 485
DB 781 GTTGGTGGAAAGCTGGACAGCTTGTGAGCATTTTGGAAAACAGAGGATCACACCA 840
QY 486 IyPProSerIeueuIyAaIaIeueuIleAaGlyAIAIAaPserAIAIyLeuGlyYr 505
DB 841 AAGCTTCTCTATTAAGCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 506 ProAaNgIyAaNgIyThraGlyAraGlyThraIeueuAaPlySerIeueuAaValaIaIy 525
DB 901 CGAAAGGTACCAAGATGGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 526 ValaAaNgIySerSerAIAeueuSerThraSerGInyAaIaThraYrThraPheThraIaThra 545
DB 961 GTGAACGATGTCAGTCTCTATTCACCAAGCAAAAGAGAGTACTGTTATGCTACT 1020
QY 546 AIAgIyIyPProLeuIyIleSerIeueuValIIPserAaPValaProAIAserThraIa 565
DB 1021 GCCGCAAGCCTTGAATAATCTCCCTGATATGATGATGATGATGATGATGATGATGAT 1080
QY 566 SerValIThraIeuaIaAaPserAaPleuValIleThraIaProAaNgIyThraRgYr 585
DB 1081 TCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 586 ValGlyAaAaPserPheSerAIAProPheAaAaenIIPaPserGlyAraAaAaValaIu 605
DB 1141 GTAGAAATGACTTACTTCTCCCATACATATGATATGATGATGATGATGATGATGATGAT 1200
QY 606 AaenValaPheIleAaenSerProGInSerGlyThraThraIleGlyAaIaIyThraa 625
DB 1201 AATGATTAATTAATGACCAACAAGCGGAGCGATACAAATTAAGTGAACAGCTTATTAAC 1260
QY 626 ValProValaIyProGInAaPserIeueuAIAIeValaAaen 639
DB 1261 GTACCGGTGACCAAGACCTTCTCTGTCATTTGTGANT 1302

RESULT 15

US-10-820-712A-21
Sequence 21, Application US/10820712A
Publication No. US2005026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mutsuyoshi
APPLICANT: Izawa, Yoshifumi
APPLICANT: Kobayashi, Tohru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820, 712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 21
LENGTH: 1302
TYPE: DNA
ORGANISM: Bacillus NCIB1289
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1302)
OTHER INFORMATION: Gene name A1, protease A1
US-10-820-712A-21

Alignment Scores:

Pred. No.:	2,08e-183	Length:	1302
Score:	2122.00 <td>Matches:</td> <td>400</td>	Matches:	400
Percent Similarity:	97.7%	Conservative:	24

Beet Local Similarity: 92.2% Mismatches: 10
 Query Match: 65.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-784-870-4 (1-639) x US-10-820-712A-21 (1-1302)

QY 206 AsnAspValAlaArgIleValIleValAlaAspValAlaGlnSerSerTyrGlyLeuTyr 225
 DB 1 AACGATGTTGCCAGAGGCATTGTAAAGCCGATGTTCCAGAGCACTATGTTTATAT 60
 QY 226 GlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSer 245
 DB 61 GGGCAAGGGCAAGTGGTGCAGTAGCGGATACCGGACTGTGATACAGGGCTTAATGACGC 120
 QY 246 SerMetHisGlnAlaPheArgGlyValIleThrAlaLeuTyrAlaLeuGlyArgThrAsn 265
 DB 121 TCGATCATGAAACGTTCCAGAGAAAGATTACCGCATATATGCCCTTGGAAACAAAC 180
 QY 266 AsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValIleGlyAsnGly 285
 DB 181 AACGGCATATCATCAACGACACGGGACGCACTTCCGATCTGTTTACGAAACGCT 240
 QY 286 AlaThrAsnIleGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSer 305
 DB 241 ACTTCGAATTAAGGATGGCTCCACAACTTAATCTTATGTTTCCATCTGTTATGACAGC 300
 QY 306 SerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSer 325
 DB 301 AATGGCGGACTTGGCGACTGCTTCATGTAAGTACATTAATTCAGCCAGGCAATATAGT 360
 QY 326 AlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAlaValAsnGlyAlaTyrThrThr 345
 DB 361 GCCGGTCCAGATTCATACGAACTATGGGAGCGCTTAAACGAGGCTTACACTACT 420
 QY 346 AspSerArgAsnValAspAspTyrValArgIleAsnAspMetThrIleLeuPheAlaIle 365
 DB 421 GATTCAGAAACGTACAGCATTAATGTTGTAATAATGATATGGCGGTTCTTTTGGCAGCG 480
 QY 366 GlyAsnGlyLeuArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaValAsnAlaIle 385
 DB 481 GGTAACGAAGGGCCGATGGCGGACATCAAGTCTCTGGTACCGGAAATGCTATC 540
 QY 386 ThrValGlyAlaThrGlnAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsn 405
 DB 541 ACAGTAGGGGCAACAGAAACCTGCCGCCAAGCTTGGATCTTAAGCTACACATCAAT 600
 QY 406 HisValAlaGlnPheSerSerArgGlyProThrIleAspGlyArgIleLeuProAspVal 425
 DB 601 CATGTAGCAAGTTTTCTCCCGGAGCTTACAAAGAGATGGAAGTATCAAAACCGGACGT 660
 QY 426 MetAlaProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerPhe 445
 DB 661 ATGGCACCGAGCAATTAATTAATGCGCAAGATCTTCTTGGCTCCGACCTCTCATTC 720
 QY 446 TrpAlaAsnHisAspSerLysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIle 465
 DB 721 TGGGCAACCATATACAGCAATATGCTTAATGGAGAACATCCATGCGGACCCGAT 780
 QY 466 ValAlaGlyAsnValAlaGlnLeuArgGlnHisPheValIleAsnArgGlyIleThrPro 485
 DB 781 GTAGCTGTGTAAGTTCAGATTAAGTGAACATTTCTATCAAAAACAGAGGAATCACTCCT 840
 QY 486 LysProSerLeuLeuValAlaIleLeuIleAlaGlyAlaAlaAspValGlyLeuGlyTyr 505
 DB 841 AAACCACTCTTGTGTAAGGAGCTTATATGCGGAGCAACTGATATCGGCTTGGCTAT 900
 QY 506 ProAsnGlyAsnGlnGlyTyrGlyArgValIleThrLeuAspLysSerLeuAsnValAlaTyr 525
 DB 901 CCGAGTGAACCAAGAGATGGGAAAGATGACATGACAAAGTCACTTAATGTAGCTTTC 960
 QY 526 ValAsnGlySerSerAlaLeuSerThrSerGlnIleValAlaThrTyrThrPheThrAlaThr 545
 DB 961 GTAAATGAACAAAGCTTTTATCTATACCAAAAGGCTACGTAATTCATTACTGCAAA 1020

QY 546 AlaGlyLysProLeuLysIleSerLeuValIleTrpSerAspAlaProAlaSerThrThrAla 565
 DB 1021 TCAAGCAACCTTTGAAGATTCATTGGTTGGTCTGAATGACCGGCAAGTACTTCCGCA 1080
 QY 566 SerValIleThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyThrArgTyr 585
 DB 1081 TCGGTTACATGGTGAATGATCTGATCTGGTATTAACAGCTCCAAATGGAACAAAGTAT 1140
 QY 586 ValGlyAsnAspPheSerAlaProPheAspAsnIleTrpAspGlyArgAsnAsnValGln 605
 DB 1141 GTTGAACAGCACTTACTGCTCCCTATGATTAATTAACGGGATGACGTAACAATGTAGAG 1200
 QY 606 AsnValPheIleAsnSerProGlnSerGlyThrTyrThrIleGlyValAlaGlyArgAsn 625
 DB 1201 AACGTGTTATCAATGCTCCGAAACGGAAACGTATACGTTGAGTACAGGCTTACAT 1260
 QY 626 ValProValGlyProGlnAsnPheSerLeuAlaIleValAsn 639
 DB 1261 GTTCCAAAGGGCCGAGCGCTTTCTTGGCTATTTGTGAC 1302

Search completed: April 8, 2006, 04:36:19
 Job time : 1356 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 8, 2006, 04:18:10 ; Search time 2708 Seconds

(without alignments)
944.001 Million cell updates/sec

Title: US-10-784-870-4

Perfect score: 3267

Sequence: 1 MRKKVFLSVLSAAALSTV.....EVQAYNPVGPQNFSLAIIV 639

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlp
-Q=abse/ABSSWB.spool/US10784870/runat.07042006.143430.27776/app.query.fasta.1
-DB=Published Applications NA New -OPMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human0.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abss05p
-USER=US10784870.qden.1.1.987.0runat.07042006.143430.27776 -NCPU=6 -ICPU=3
-NO MAMP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOB=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:*
1: /SIDS5/ptodata/1/pubpna/US08 NEW PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06 NEW PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07 NEW PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09 NEW PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09 NEW PUB.seq:*
7: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
8: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
9: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
10: /SIDS5/ptodata/1/pubpna/US10 NEW PUB.seq:*
11: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
12: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
13: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
14: /SIDS5/ptodata/1/pubpna/US11 NEW PUB.seq:*
15: /SIDS5/ptodata/1/pubpna/US60 NEW PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315	9.6	3625	8 US-10-510-386-27	Sequence 27, Appl
2	315	9.6	4041	8 US-10-510-386-199	Sequence 199, App

	3	292.5	9.0	3405	8	US-10-510-386-1	Sequence 1, Appl
	4	260.5	8.0	1140	14 <th>US-11-156-062-22</th> <th>Sequence 22, Appl</th>	US-11-156-062-22	Sequence 22, Appl
	5	245	7.5	1495	14 <th>US-11-020-602-1</th> <th>Sequence 1, Appl</th>	US-11-020-602-1	Sequence 1, Appl
	6	244	7.5	4216	14 <th>US-11-136-527-605</th> <th>Sequence 605, Appl</th>	US-11-136-527-605	Sequence 605, Appl
	7	243	7.4	4338	14 <th>US-11-122-329-113</th> <th>Sequence 113, Appl</th>	US-11-122-329-113	Sequence 113, Appl
	8	242.5	7.4	1180	11 <th>US-11-047-380-4</th> <th>Sequence 4, Appl</th>	US-11-047-380-4	Sequence 4, Appl
	9	242	7.4	4199	14 <th>US-11-128-061-679</th> <th>Sequence 679, Appl</th>	US-11-128-061-679	Sequence 679, Appl
	10	242	7.4	4199	14 <th>US-11-128-061-679</th> <th>Sequence 679, Appl</th>	US-11-128-061-679	Sequence 679, Appl
	11	238.5	7.3	5236	8 <th>US-10-510-386-217</th> <th>Sequence 217, Appl</th>	US-10-510-386-217	Sequence 217, Appl
	12	226.5	6.9	824	14 <th>US-11-156-062-11</th> <th>Sequence 13, Appl</th>	US-11-156-062-11	Sequence 13, Appl
	13	226.5	6.9	825	14 <th>US-11-065-943-59</th> <th>Sequence 59, Appl</th>	US-11-065-943-59	Sequence 59, Appl
	14	225.5	6.9	824	14 <th>US-11-156-062-11</th> <th>Sequence 11, Appl</th>	US-11-156-062-11	Sequence 11, Appl
	15	223	6.8 <td>2760</td> <td>13<th>US-11-008-331-1</th><th>Sequence 1, Appl</th></td>	2760	13 <th>US-11-008-331-1</th> <th>Sequence 1, Appl</th>	US-11-008-331-1	Sequence 1, Appl
	16	222.5	6.8 <td>824</td> <td>14<th>US-11-156-062-7</th><th>Sequence 7, Appl</th></td>	824	14 <th>US-11-156-062-7</th> <th>Sequence 7, Appl</th>	US-11-156-062-7	Sequence 7, Appl
	17	221.5	6.8 <td>824</td> <td>14<th>US-11-156-062-3</th><th>Sequence 3, Appl</th></td>	824	14 <th>US-11-156-062-3</th> <th>Sequence 3, Appl</th>	US-11-156-062-3	Sequence 3, Appl
	18	220.5	6.7 <td>822</td> <td>14<th>US-11-156-062-9</th><th>Sequence 9, Appl</th></td>	822	14 <th>US-11-156-062-9</th> <th>Sequence 9, Appl</th>	US-11-156-062-9	Sequence 9, Appl
	19	219.5	6.7 <td>824</td> <td>14<th>US-11-156-062-5</th><th>Sequence 5, Appl</th></td>	824	14 <th>US-11-156-062-5</th> <th>Sequence 5, Appl</th>	US-11-156-062-5	Sequence 5, Appl
	20	217	6.6 <td>1437</td> <td>9<th>US-10-932-182A-2353</th><th>Sequence 2353, Ap</th></td>	1437	9 <th>US-10-932-182A-2353</th> <th>Sequence 2353, Ap</th>	US-10-932-182A-2353	Sequence 2353, Ap
	21	217	6.6 <td>1437</td> <td>9<th>US-10-932-182A-2353</th><th>Sequence 2353, Ap</th></td>	1437	9 <th>US-10-932-182A-2353</th> <th>Sequence 2353, Ap</th>	US-10-932-182A-2353	Sequence 2353, Ap
	22	216.5	6.6 <td>1437</td> <td>9<th>US-10-932-182A-2353</th><th>Sequence 2353, Ap</th></td>	1437	9 <th>US-10-932-182A-2353</th> <th>Sequence 2353, Ap</th>	US-10-932-182A-2353	Sequence 2353, Ap
	23	216.5	6.6 <td>824</td> <td>14<th>US-11-156-062-15</th><th>Sequence 15, Appl</th></td>	824	14 <th>US-11-156-062-15</th> <th>Sequence 15, Appl</th>	US-11-156-062-15	Sequence 15, Appl
	24	215.5	6.6 <td>824</td> <td>14<th>US-11-156-062-17</th><th>Sequence 17, Appl</th></td>	824	14 <th>US-11-156-062-17</th> <th>Sequence 17, Appl</th>	US-11-156-062-17	Sequence 17, Appl
	25	211.5	6.5 <td>4944</td> <td>14<th>US-11-052-554A-643</th><th>Sequence 643, Appl</th></td>	4944	14 <th>US-11-052-554A-643</th> <th>Sequence 643, Appl</th>	US-11-052-554A-643	Sequence 643, Appl
	26	205	6.3 <td>2772</td> <td>8<th>US-10-510-386-21</th><th>Sequence 21, Appl</th></td>	2772	8 <th>US-10-510-386-21</th> <th>Sequence 21, Appl</th>	US-10-510-386-21	Sequence 21, Appl
	27	197.5	6.0 <td>1476</td> <td>9<th>US-10-932-182A-75743</th><th>Sequence 75743, A</th></td>	1476	9 <th>US-10-932-182A-75743</th> <th>Sequence 75743, A</th>	US-10-932-182A-75743	Sequence 75743, A
	28	197.5	6.0 <td>1476</td> <td>9<th>US-10-932-182A-75743</th><th>Sequence 75743, A</th></td>	1476	9 <th>US-10-932-182A-75743</th> <th>Sequence 75743, A</th>	US-10-932-182A-75743	Sequence 75743, A
	29	197.5	6.0 <td>2566</td> <td>11<th>US-11-096-568A-24711</th><th>Sequence 24711, A</th></td>	2566	11 <th>US-11-096-568A-24711</th> <th>Sequence 24711, A</th>	US-11-096-568A-24711	Sequence 24711, A
	30	190	5.8 <td>1908</td> <td>9<th>US-10-932-182A-76207</th><th>Sequence 76207, A</th></td>	1908	9 <th>US-10-932-182A-76207</th> <th>Sequence 76207, A</th>	US-10-932-182A-76207	Sequence 76207, A
	31	190	5.8 <td>1815</td> <td>9<th>US-10-932-182A-76207</th><th>Sequence 76207, A</th></td>	1815	9 <th>US-10-932-182A-76207</th> <th>Sequence 76207, A</th>	US-10-932-182A-76207	Sequence 76207, A
	32	189	5.8 <td>1815</td> <td>9<th>US-10-932-182A-76207</th><th>Sequence 76207, A</th></td>	1815	9 <th>US-10-932-182A-76207</th> <th>Sequence 76207, A</th>	US-10-932-182A-76207	Sequence 76207, A
	33	189	5.8 <td>1815</td> <td>9<th>US-10-932-182A-76207</th><th>Sequence 76207, A</th></td>	1815	9 <th>US-10-932-182A-76207</th> <th>Sequence 76207, A</th>	US-10-932-182A-76207	Sequence 76207, A
	34	182	5.6 <td>4856</td> <td>14<th>US-11-136-527-3850</th><th>Sequence 659, Appl</th></td>	4856	14 <th>US-11-136-527-3850</th> <th>Sequence 659, Appl</th>	US-11-136-527-3850	Sequence 659, Appl
	35	175.5	5.4 <td>2251</td> <td>14<th>US-11-128-061-10</th><th>Sequence 3850, Appl</th></td>	2251	14 <th>US-11-128-061-10</th> <th>Sequence 3850, Appl</th>	US-11-128-061-10	Sequence 3850, Appl
	36	175.5	5.4 <td>2251</td> <td>14<th>US-11-128-061-10</th><th>Sequence 10, Appl</th></td>	2251	14 <th>US-11-128-061-10</th> <th>Sequence 10, Appl</th>	US-11-128-061-10	Sequence 10, Appl
	37	175.5	5.4 <td>2388</td> <td>14<th>US-11-128-061-1086</th><th>Sequence 1, Appl</th></td>	2388	14 <th>US-11-128-061-1086</th> <th>Sequence 1, Appl</th>	US-11-128-061-1086	Sequence 1, Appl
	38	175.5	5.4 <td>2543</td> <td>11<th>US-11-096-568A-11250</th><th>Sequence 31250, A</th></td>	2543	11 <th>US-11-096-568A-11250</th> <th>Sequence 31250, A</th>	US-11-096-568A-11250	Sequence 31250, A
	39	175.5	5.4 <td>4190</td> <td>8<th>US-10-821-234-324</th><th>Sequence 324, Appl</th></td>	4190	8 <th>US-10-821-234-324</th> <th>Sequence 324, Appl</th>	US-10-821-234-324	Sequence 324, Appl
	40	171.5	5.2 <td>4626</td> <td>14<th>US-11-000-688-445</th><th>Sequence 445, Appl</th></td>	4626	14 <th>US-11-000-688-445</th> <th>Sequence 445, Appl</th>	US-11-000-688-445	Sequence 445, Appl
	41	168	5.1 <td>2611</td> <td>11<th>US-11-096-568A-11860</th><th>Sequence 31860, A</th></td>	2611	11 <th>US-11-096-568A-11860</th> <th>Sequence 31860, A</th>	US-11-096-568A-11860	Sequence 31860, A
	42	166.5	5.1 <td>2527</td> <td>14<th>US-11-128-061-1086</th><th>Sequence 1086, Ap</th></td>	2527	14 <th>US-11-128-061-1086</th> <th>Sequence 1086, Ap</th>	US-11-128-061-1086	Sequence 1086, Ap
	43	166.5	5.1 <td>2527</td> <td>14<th>US-11-128-061-1086</th><th>Sequence 1086, Ap</th></td>	2527	14 <th>US-11-128-061-1086</th> <th>Sequence 1086, Ap</th>	US-11-128-061-1086	Sequence 1086, Ap
	44	161	4.9 <td>2706</td> <td>11<th>US-11-096-568A-17893</th><th>Sequence 17893, A</th></td>	2706	11 <th>US-11-096-568A-17893</th> <th>Sequence 17893, A</th>	US-11-096-568A-17893	Sequence 17893, A
	45	158.5	4.9 <td>2537</td> <td>11<th>US-11-096-568A-23981</th><th>Sequence 23981, A</th></td>	2537	11 <th>US-11-096-568A-23981</th> <th>Sequence 23981, A</th>	US-11-096-568A-23981	Sequence 23981, A

ALIGNMENTS

RESULT 1
US-10-510-386-27
Sequence 27, Application US/10510386
Publication No. US20050244922A1
GENERAL INFORMATION:
APPLICANT: Andersen, Jens Tonne
APPLICANT: Clausen, Ib Groth
APPLICANT: Jorgensen, Steen Troels
APPLICANT: Olsen, Peter Bjarke
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: Improved Bacillus Host Cell
FILE REFERENCE: 10294.204-US
CURRENT APPLICATION NUMBER: US/10/510,386
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 248
SOFTWARE: PatentIn version 3.3
SEQ ID NO 27
LENGTH: 3625
TYPE: DNA
ORGANISM: Bacillus licheniformis
FEATURE:
NAME/KEY: CDS
LOCATION: (501)..(3122)
US-10-510-386-27

Alignment Scores:

Pred. No.:	9.06e-22	Length:	3625
Score:	3.5.00	Matches:	161
Percent Similarity:	41.4%	Conservative:	92
Best Local Similarity:	26.4%	Mismatches:	234
Query Match:	9.6%	Indels:	125
DB:	8	Gaps:	26

US-10-784-870-4 (1-639) x US-10-510-386-27 (1-3625)

QY	2	ArglyvslvlyValPheLeuserValIeu-serIalalalaleuserThValI	21
Db	1106	AGGAAAAGAGCTGTTTC---CTTACTATTAAGAGCTTCATTCATGCGCAAAAT	1162
QY	21	aleuabnaaProSerIalglAspIalargThPheaBpIeuaep-----	36
Db	1163	GCTCTTTAACAATCGATGAGAGACAGGTGTCTACAAAGACCTGGTGCAGCTGAAGCCGT	1222
QY	37	-PheIyagIylIegIlnThrThraBpValserGlyPheSer---LyvGlnArgGlnTh	55
Db	1223	GTTTGCAAGATTCGCTTAAACCGACAAAGTGAGCCCAATCACTCGATTACGAATGATGATCA	1282
QY	55	rgIyAlalalalPheIeValIgluserGluabnaValIyIeueuIyrgIyIeueIy	75
Db	1283	AAAAGCCATCGCTGCTGTACGAACACCTCGCGGTCGTTCCGAGCCGTTGAAAAA	1342
QY	75	slYleuIgluThrValProalaaBnaIyIeueThslIeValIglPheaangIy----	93
Db	1343	ACAGCTGGATCAAGTCGCGAAAGAC-----ATCGGATTCGAACATTAACGAGCAGCA	1396
QY	94	-----ProIleuIgluIuThrLyvGlnIyIeueIgluThrThrgIyAlalyIleLe	111
Db	1397	AGTATCGGCTTGCTGTGAAAAAGCGGGATGGCGACATTCATCAACGCG-----	1448
QY	111	uaBpTyrlleProaBpTyrlalTyrlIeValIgluIyrgIu-----GlyAs	126
Db	1449	-----CCGAAACCGTTACATGTGAATTAAGAAAGCGCAAAAAACCGGAGTC	1498
QY	126	pvalIgluserIyBvalArgserIleGluHleValIgluserValIguPro-----	142
Db	1499	TTTCAAACTTAAGGCCAATCA---TCCGGGCTCAGGCACTTAAGACCCCTCGGTAAAG	1555
QY	143	-----TyrlEuProLyTyrlIyIleBpProGlnIeupheth	155
Db	1556	CAAAACGCGCATTTAAAGATATGTACGTTGTGAAATGAAGAAAGCGGTTCTCCGAGTT	1615
QY	155	rlYvGlyAlaserThrieuValIyBAlIeuaIaleuBpThrxysGlnaBnaIyrgI	175
Db	1616	CAAAACGCGCGCAAGCAATATCAAGCGGCGCGCC-----TCAAGATCGCCAAAT	1666
QY	175	uvalIgluBnaIyrgIylIegIuGluIlealIglIyThValIa--SerBnaBpValH	194
Db	1667	GCGTGAAGTGAATTCGTGCAACAGGTTCAACCAATTAAGACACTGTCAAGACACCCA	1726
QY	194	slTyrlIerThralalyBpGluTyrlYlvBvalIeCaaBnaBpValIalargIylIeValIy	214
Db	1727	ATAT-----CCATATCAATGATGCTCAAAAATTAACGCAAAAACGCTGCGGAA	1777
QY	214	BAlaBpValIaleGlnBserTyrgIyIeueTyrgIyGlnIyGln-----	229
Db	1778	TGCGACATCAATTTGAACAGCTTCAGAAAGCTGATGAAGCAAAAGCTGAAGAATAC	1837
QY	230	-IleValIalValAlaBpThrxIyIeuaBpThrxIyArGynaBpBserBermethIsgI	249
Db	1838	AGTAATGCGCGTGGTGAACACAGCGCTTGATCAACCTTGGGATTTAAGC-----	1889
QY	249	uAlaBpAerGlyIyIeIe-----ThralaIeueTyrlaIeueIyIyrgThraBna	266
Db	1890	-----GGCAGCGTCAAAAAGACGAAGGCTATACTATGTGCGCGCACGGCGGA	1939
QY	266	nAlaBnaBpThraBnIyHIsGlyThrxIyValIalglYserValIeueIy-----	283
Db	1940	TGCGATGAATGAACATGCGCACCGGACACACGTGTGCGGATCACTTCAGCGCGCAGAA	1999

QY	284	AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleu	303
Db	2000	CACACATTTTGGATGGCGGGAATCAATGCTTATTCGCAAAATTCGCTCGAAGTGC	205
QY	303	LaPseSerSergLyGlyLeuGlyGly-LeuProSerAsnLeuGlnThrLeuPheSerGln	323
Db	2060	GGATTCCTCAGGC-----AGTCGAGATACGAAACAGATTGCCAACCGGATCATCTAT-G	211
QY	323	LaPheSerAlaGlyAlaArgIleHsrAsnSerTrpGlyAlaAlaValaAsnGlyAlaT	343
Db	2113	CGCGCGAACCAACGGTCCAAAGATCAATTTAAAGTCTGGC-----GGGCCAT	216
QY	343	YrThrThraPseSerArgAsnValaPaPeryrValArgLyAsnAspMetThrIleLeuP	363
Db	2161	ACAGCGGGGTGATGGAATATGCGCTTAAATATAGCGGCTTCTTAAATATGAGAGATCGTGG	222
QY	363	heAlaAlaGlyAsnGlnArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLys	383
Db	2221	CGCGCACCGGAATATAC-----GGAGTATCGGAGATTTCTTACCTTGATCTTGGAAAT	227
QY	383	snAlaIleThrValGlyAlaThrGlyAsnLeuArgProSerPheGlySeryrAlaAsp	403
Db	2275	ATAGCCTTACGTGGGGGAGCAATATATCTC-----	230
QY	403	snIleAsnHisValAlaGlnPheSerSeraGlyProThrLysAspGlyArgIleLysP	423
Db	2307	-----GATTCGTCTCGAGACTACTCAATTTATGGAAGAGTCTC-----	234
QY	423	roAspValMetAlaProGlyThrTrpIleLeuSerAlaArgSerSerLeuAlaProAsp	443
Db	2346	--GATATGTGTGCGCGGGAACCGATATTT-----CCAGGCTCGTTCCGGACG	239
QY	443	erSerPheTrpAlaAsnHisAspSerLysrAlaLysrMetGlyGlyThrSerMetAlaT	463
Db	2392	GGAAAT-----GTCACCTATATAGCGGAAATCATGATGGCGG	242
QY	463	hrProIleValAlaGlyAsnValAlaGlnLeuAspGlyHisPheValLysAsnArgLyI	483
Db	2428	CGCGCGACGTCGACGCTGCACAGACCTTCTT-----TTGTCACGAATC	247
QY	483	LeThrProLysProSerLeuLeuLysAlaAlaLeuIleAlaGlyAlaAlaAspValGlyL	503
Db	2473	CGTCTTGAAACCAAGCAAAATCGCAACCTATATGACCGAGACGACAGCATATGTGCAT	253
QY	503	eu-----GlyTrpAsnGlyAsn	509
Db	2533	TTGAAGACGACGATATATCCAAACCGGATTTATGACCTGATATAGAACCGGCTGCACAA	259
QY	510	-----GlnGlyTrpGlyArgValThrLeuAspLysSerLeuAsnV	523
Db	2593	TTCCCGGATATGACTTCGTCTCCGGGTGGGAGAGGCTGAATGTTTTCATCGACCGACG	265
QY	523	AlaIaLysrValAsnGlySerSerAlaLeuSerThrSerGlnLysAlaThrTrpThrPheT	543
Db	2653	TTTTTGAAGCTAAACATGAAGTTCAATCCGTTTTTAACCGGCATACGCGAGTACAGGCA	2712
QY	543	hrAlaThrAlaGlyLysProLeuLysIle	552
Db	2713	CAGCCAAAGCGGTGTGACGCTCAAAATC	2741
RESULT 2			
US-10-510-386-199			
; Sequence 199, Application US/10510386			
; Publication No. US20050244922A1			
; GENERAL INFORMATION:			
; APPLICANT: Andersen, Jens Tonne			
; APPLICANT: Clausen, Ib Groth			
; APPLICANT: Jorgensen, Steen Troels			
; APPLICANT: Olsen, Peter Bjarke			
; APPLICANT: Rasmussen, Michael Dolberg			
; TITLE OR INVENTION: Improved Bacillus Host Cell			
; FILE REFERENCE: 10294.204-US			
; CURRENT APPLICATION NUMBER: US/10/510.386			

CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 248
SOFTWARE: PatentIn version 3.3
SEQ ID NO 199
LENGTH: 4041
TYPE: DNA
ORGANISM: Bacillus licheniformis
FEATURE:
NAME/KEY: CDS
LOCATION: (501) .. (3641)
US-10-510-386-199

Alignment Scores:
Pred. No.: 1,03e-21 Length: 4041
Score: 315.00 Matches: 161
Percent Similarity: 41.4% Conservative: 92
Best Local Similarity: 26.4% Mismatches: 234
Query Match: 9.6% Indels: 125
Gaps: 26

US-10-784-870-4 (1-639) x US-10-510-386-199 (1-4041)

QY 2 ArglyLeuValPheLeuSerValLeu-SerAlaAlaAlaLeuSerThrValAl 21
DB 1106 AGGAAAGAGCTGCTTC---CCTTACTATAAGAGCTCCATTCAGCGCAAAAT 1162
QY 21 AlaLeuAsnProSerAlaGlySerAlaArgThrPheAspLeuAsp----- 36
DB 1163 GCTCTTAAACAAATTCAGTAGAGACAGTCTTCAAAAGACTGTGACGTGAAGCGCT 1222
QY 37 -PheUsglyLeuThrThrThrAspValSerGlyPheSer---LysGlnArgGlnTh 55
DB 1223 GTTTCAGATGTCGCTTAAAGACGACAGTAGAGCGCATCTCAATTCAGATGATGATCA 1282
QY 55 rglYalAlaAlaPheLeuValGluSerGluAsnValLysLeuLeuUsglyLeuLeu 75
DB 1283 AAAAGGATCAGTCGCTGTACGAAACAGCTCGCGGCTCCGAGCGCTTGAAGAAA 1342
QY 75 elyLeuGluThrValPheAlaAsnAsnLysLeuHisLeuAlaGlnPheAsnGly----- 93
DB 1343 ACAGCTGATCAGTGCAGCGCAAGAC-----ATCGCATTCGAAACAATTACAGGACGAA 1396
QY 94 -----ProIleuGluGluThrLysGlnLysGluThrThrGlyAlaLysIlele 111
DB 1397 AGTATCGGCTGCTGTGAAAGACCGGAGATGCGACAGCTTCAACGCG----- 1448
QY 111 uAspTyrlleProAspTyrlaTyrlleValGluTyrglu-----GlyAs 126
DB 1449 -----CCGAAACCGCTTACATCTGTAATAATTGAAAGACGCAAAACCGGATC 1498
QY 126 pValGlnSerLysValArgSerIleGluHisValGluSerValGluPro----- 142
DB 1499 TTTCAATCTAAAGCCCAATCA---TCCGCGCTCAGGCACTAGACCCCTCGTAAAG 1555
QY 143 -----TyrlleuProLysTyrlleAspProGlnLeuAspTherm 155
DB 1556 CAAAGCGCATTTAAAGATATGATGCTGTGGAATGAGAAAGCCGTTCTTCGAGATT 1615
QY 155 rlyGlyAlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysG 175
DB 1616 CAAAGCGCGGCAAAACAAATTCAGCGCGAGCC-----TCCAGATGCCCAAGAT 1666
QY 175 uValGlnLeuAspGlyIleGluGluIleAlaGlnTyrgluAla---SerAsnAspValH 194
DB 1667 GCCTGAAGTGGAATTCGTGCAGACAGGTTGAGCAATGAAAGCACTGTCAAGACACCCA 1726
QY 194 sTyrlleThrAlaLysProGluTyrlLysValMetAsnAspValAlaArgGlyIleVal 214
DB 1727 ATRATY-----CAATATCAATGTGCTCAAAATAACCGCAAAACCGTCTCGAA 1777
QY 214 sAlaAspValAlaGlnSerSerTyrglyLeuTyrglyGlnGlyGln----- 229
DB 1778 TGCTGACATATCAATTTGAACAGCTTCAAGAGCTGAGAAAGCAAAAGCTGAAGATAC 1837

QY 230 -IleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHisG 249
DB 1838 AGTAATCGCGCTGTGACACAGCGTATATCATACCTTCGCGAATTTAAGC----- 1889
QY 249 uAlaPheArgLysIle-----ThrAlaLeuTyrlAlaLeuGlyArgThrAsnAs 266
DB 1890 -----GCGAGCGTCAAAAAGACGAAGCTTAACTATGTCGCGCGCACCGCGGA 1939
QY 266 nAlaAsnAspThrAsnGlyHisIleGlyThrHisValAlaGlySerValLeuGly----- 283
DB 1940 TGGCATGATGACAAATGCGCACGCGCACACGTCGACGATCATTCACCGCGCAAGA 1999
QY 284 -AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGlnSerIle 303
DB 2000 CAACCAATTTTCATGCGCGGAATCAATGCTTATGCTCAAAATTCCTGTCGAAAGCT 2059
QY 303 AspSerSerGlyGlyLeuGlyGlyLysProSerAsnLeuGlnThrLeuPheSerGln 323
DB 2060 GATTTCTTACGC-----AGTGAATACGAAACGATTCGAAACGCGCATCTAT-G 2112
QY 323 lAspSerAlaGlyAlaArgIleHisThrAsnSerTrpGlyAlaAlaValAsnGlyAla 343
DB 2113 CCGCGACACACGCTGCAAAAGTCATCAATTTAAGTCTTGCG-----GGGCGAT 2160
QY 343 yThrThrAspSerArgAsnValAspAspTyrlValArgLysAsnAspMetThrIleLeu 363
DB 2161 ACAGCGGCTGATGGAATATGCGCTTAAATATCGCGCTTCAAAATGTCGACATGCTG 2220
QY 363 heAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThrAlaLys 383
DB 2221 CCGCACCGGAATATGAC-----GAGATATCGAAGATTTCTTACCTGTCATCTTCGAAT 2274
QY 383 sAlaIleThrValGlyAlaArgGluAsnLeuArgProSerPheGlySerTyrlAsp 403
DB 2275 ATACGCTTCAAGTCGCGGACCAATATATCTC----- 2306
QY 403 snlLeuAsnHisValAlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLys 423
DB 2307 -----GATCTTGTCTGAGATCTCCAAATATGAAAGATCTC----- 2345
QY 423 roAspValMetAlaProGlyThrTyrlleLeuSerAlaArgSerSerLeuAlaProAsp 443
DB 2346 --GATATGCTGCGCGGACCGAATAT-----CCAGCTTCGTTCCGAGC 2391
QY 443 eSerPheThrAlaAsnHisAspSerLysTyrlaTyrlMetGlyGlyThrSerMetAla 463
DB 2392 GGAT-----GTCACTATATGAGCGGAACATGATGCGCG 2427
QY 463 h-ProIleValAlaGlyAsnValAlaGlnLeuArgGluHisPheValLysAsnArgGly 483
DB 2428 CGCGCACGTCGACAGTCGACGACGACTTCT-----TTGTCAACAATC 2472
QY 483 lAspProLysProSerLeuLeuLysValAlaLeuAlaLeuAlaGlyAlaAlaAspVal 503
DB 2473 CGCTTGAACCAAAACCAATTCGCAAGCTTATGACGAGACGACGACGATGTGCAT 2532
QY 503 eu-----GlyTyrlProAsnGlyAsn----- 509
DB 2533 TTGAAGACAGATATATCAAAACCGGATTAATGACTGATATGAAACCGGCTGCACAA 2592
QY 510 -----GlnGlyTyrlGlyArgValThrLeuAspLysSerLeuAsn 523
DB 2593 TTCCCGGATATGACTTCTGCTCGCGGTGGGAAAGCTGATATTTTTCATGACGCGAG 2652
QY 523 sAlaIleTyrlValAsnGluSerSerAlaLeuSerThrSerGlnLysAlaThrTyrlThrPhe 543
DB 2653 TTTTGAAGCTGAACATGAAGTTTCATCCCGTTTAAACCGCATACGCGAGTGAAGCA 2712
QY 543 hAlaThrAlaGlyLysProLeuLysIle 552
DB 2713 CAGCCAAAGCGGTGACGTCGCAAAATC 2741

```
RESULT 3
US-10-510-386-1
/ Sequence 1, Application US/10510386
/ Publication No. US20050244922A1
/ GENERAL INFORMATION:
/ APPLICANT: Andersen, Jens Tonne
/ APPLICANT: Andersen, Ib Groth
/ APPLICANT: Clausen, Steen Troels
/ APPLICANT: Jorgensen, Michael Dolberg
/ APPLICANT: Rasmussen, Michael Dolberg
/ TITLE OF INVENTION: Improved Bacillus Host Cell
/ FILE REFERENCE: 10294.204-US
/ CURRENT APPLICATION NUMBER: US/10/510,386
/ NUMBER OF SEQ ID NOS: 248
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 1
/ LENGTH: 3405
/ TYPE: DNA
/ ORGANISM: Bacillus licheniformis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (501)..(2906)
/ US-10-510-386-1

Alignment Scores: %
Pred. No.: 1,77e-19 Length: 3405
Score: 292.50 Matches: 166
Percent Similarity: 34.2% Conservative: 84
Beech Local Similarity: 22.7% Mismatches: 219
Query Match: 9.0% Indels: 263
DB: 8 Gaps: 33

US-10-784-870-4 (1-639) x US-10-510-386-1 (1-3405)

QY 126 APPVALGINSErLyValrSerIleGIuHsValGluSerValGluProTyLeuPro 145
DB 645 GATGTGACGCTGTGATTAACAGACGACGCGTTATCGTGAACTGAAAGAAAGCGCTTGGC 704
QY 146 LysTyLyVLIeAsPProGIuLeuHeTherLyVSGIYAlaSerThLeuValIyAlaLeu 165
DB 705 GAAGCAAAAGCCGACGAGAAACAAACAAAGGCTTCTTAACAAACAGCTTGAAACAA 764
QY 166 AlaLeuAsPThLySGInAsn-----AsnLySGIuValGIn-----LeuArgGIY 180
DB 765 GCGCTGAAACACTTAATAAAAGCGAAAGTAACCCGAAATACGACCGTATTTCCGGC 824
QY 181 Ile-----GluGluIleAlaGInTyValAlaSerAsnAsPValHis 194
DB 825 TTTTCTATGAAACCTCCGCGCAGTGAATTCCAAAGCTGCTCGCGTCAAGAAATTA 884
QY 195 TyrIle-----ThraIaLysProGIu----- 201
DB 885 GCGGTTTATCCAAACGCACTTAACAACCTGACAGCGTAAAGAAAGAAAGACGTGACACTT 944
QY 202 -----TyLyVaIeMeTAsnAsPValAlaArgLyIleValIyAlaAsP 216
DB 945 GCAGACAGACGCATTATTCACAGATG---GATTAAGATGCGCCCTTCATCGAGCGGAT 1001
QY 217 ValAlaGInSerSerTyLyGIYLeuTyGIYGIYGIYGIYGIYGIYGIYGIYGIYGIY 236
DB 1002 CAGCGATGAAATCG---GGCTATACCGGTAAAGCATTAAGGTAGCGGTATTGATGACG 1058
QY 237 GIYLeuAsPThrGIYArGAsnAsP---SerSerMetHisGluAlaPheArgGIYlys--- 254
DB 1059 GCGCGGATTAACCCATCCGACCTTAAGAAATATTTCCGCCCATATTAAGGCGGTANTGAC 1118
QY 255 ---IleThraIaLeuTyValaLeuGIYArGThraAsnAlaAsnAsPThraSngIY--- 272
DB 1119 TTTGTGATTAATGATTAATGATTCACAGAAACACCGACAGGCGACCGCGGCGGAGCG 1178
QY 273 -----HisGIYThraIeValAlaGIYSerValLeuGIYAsnGIYAlaThraAsnLySGIY 290
DB 273 -----HisGIYThraIeValAlaGIYSerValLeuGIYAsnGIYAlaThraAsnLySGIY 290
```

```
DB 1179 ACTGATTCACGGCACTCATGTTGCAGAAACAATCGCCCGAAGCGACAGATT---AAAGGC 1235
QY 291 MetAlaProGIuAlaAsnLeuValPheGInSerIleMetAsPserSerGIYGIYLeuGIY 310
DB 1236 GTGGCGCTGAAGCAACGCTTCTTGTCTTAACCGCGTCTCGGCCCGGGCGG-----TCA 1289
QY 311 GIYLeuProSerAsnLeuGInThraLeuPheSerGIYAlaPheSerAlaGIYAlaArgIle 330
DB 1290 GGCACAAACCGAAGATGTATGTCGGGCATTTGAAACCGCTTGACAGCGGACAAAGATG 1349
QY 331 HisThraAsnSerTrpGIYAlaAlaValaSngIYAla---TyThThraPserTrpAsn 349
DB 1350 ATGAACCTCTCATTTGGGAAATTCGTTGAACGCTTGATTTTGCACAAAGCATC----- 1403
QY 350 ValaAsPTrYValaArgLyAsnAsPMeThrIleLeuPheAlaAlaGIYAsnGIYArg 369
DB 1404 GCACGTGACGTGGCGCATGCGTGAAGGGGTGTGCGCGTTACGTCAACAGTAAACGCGGA 1463
QY 370 ProAsnGIYGIYThraIleSerAlaProGIYThraIyAsnAlaIleThraValGIYAla 389
DB 1464 CCGGAAACCTGACGCGTGGCTCGCCGGAAACATCAAGGTTGCGATTCCGTCGCGCA 1523
QY 390 ThrGIu-----AsnLeuArgProSerPheGIYSer----- 399
DB 1524 TCACAGCTTCGGTAAATGAATATTCGTTACGCTTCTTGTATTCATCAGCAAGGTG 1583
QY 399 ----- 399
DB 1584 ATGGATATCAAGAGAAAGAAAGATCTTGAAACGATTGAACGTTCAAGAGTGAACCTGTT 1643
QY 399 ----- 399
DB 1644 GAAGCCGCGCTCGACAAAGCTGACGATTTCTCAGGCAAAAGATGAAAGAAAGTTGCG 1703
QY 400 -----TyraIaAsPAsnIleAsnHis----- 406
DB 1704 GTTATTCAGAGAGCGGTCAATCATTTGTTGATTAAGCTGAAATTCGCAAAAACGCGGC 1763
QY 406 ----- 406
DB 1764 GCTATCGAGCGGTGATTTACATATATGCAACAGGGAATTCGAGCAACGTCATGCGGG 1823
QY 406 ----- 406
DB 1824 ATGGCCGTACGACCGTCAAAATTTCTCAAAAGAGAGGCGCAAAAGCTTTGTTCAACAGATC 1883
QY 406 ----- 406
DB 1884 AAAGAAAGGAAACACTCGGTTGCTTTCTTCAATTTGCAAAAAGCTGGGGAAACA 1943
QY 407 ValAlaGInPheSerSerArgGIYProThLyAsP---GIYArgIleYsProAsPVal 425
DB 1944 ATTGCTCATCTTCGTCGCGCGGCTGTGATGATCATGATGATTAACCTGACGTT 2003
QY 426 MetAlaProGIYThrTyLyIleLeuSerAlaArgSerSerLeuAlaProAsPserSerPhe 445
DB 2004 TCAGCGCAGGCGTCAACATCTGCAGACCATTCGAC----- 2042
QY 446 TrpAlaAsnHisAsPserLyS-----TyraIaTyMetGIYGIYThrSerMet 461
DB 2043 -----CACGATCGGAAACCCGTAACGTTACGTTCAAAACAGGAAACAGCATG 2093
QY 462 AlaThraProIleValAlaGIYAsnValAlaGInLeuArgGIYHisPheValIyAsnArg 481
DB 2094 GCTTCCCGCATGTGCGGGAAACGCGCATTTAAAGAG----- 2135
QY 482 GIYIleThraPro-----LysProSerLeuLeuValAlaIleLeuIleAlaIle 499
DB 2136 ---GCCAAACCGAATTGAGCGCTTGAGCAAAATCAAAAGCGTACTGATGAATACGCGGAA 2192
QY 500 AsPVal-----GIYLeuGIYTyYProAsnGIYAsnGIYGIYTrpGIYArgVal 515
DB 2193 AAACGTGCGATGAAATGGAAGCGCTTCCTCAAAATGCGAAGAGGCGGCGAGCATC 2252
```

```

QY 516 ThrLeuApbysSerLeuAenValAlaTyrValaenGluSerSerAlaLeuSerThrSer 535
DB 2253 CGCATCATGAGGCCCTT-----AAACCTTCATCGATTGTAACGCCG 2294
QY 536 GlnLysAlaThrTyr--ThrpheThrAlaThrAlaGlyLysProLeuLys----- 551
DB 2295 GGGAGCCATTCATACGGAACATCTCTGAAGAACAGAGGAAACAGACGAAACAAAGCG 2354
QY 551 ----- 551
DB 2355 TTCACGATTGAAACCTTCTTCAACAGAAAGCTATCAGCTCGAATACCTCTTAA 2414
QY 552 -----11eSerLeuValTrpSerAsp-----AlaProAlaSerThrThr-- 564
DB 2415 GGAACGGGCATCCGCGTATCAGAAACGAGTCGTGTACCGCCCAATCAACAGCGT 2474
QY 565 -----AlaSerValThrLeuValAaenApbLeuApbLeuValIleThrAlaProAenGly 582
DB 2475 AAGGAGCGGCAAAAGTAACCGTCATTCGCCGAAACGAAAGCAGACATATGAAGGC 2534
QY 583 ThrArgTyrValGlyAaenApbSerAlaProPheAaenAaenTrpAspGlyArgAaen 602
DB 2535 ACCGTTTACATCCGTGAA-----GACGGAAGAA 2564
QY 603 AenValGlu-----AenValPheIleAaenSerPro----- 612
DB 2565 GTCCGCGAATCCCGCTCCTATGATGCTCAAAAGCAGACATACCGCGCTCACATCC 2624
QY 613 -----GlnSerGlyThrTyrThrIleGluValGlnAlaTyrAaen 635
DB 2625 GTACAGCTTGAACCGGAGACGAGGCGCTTACCGATCGA-----GCCTAC-- 2675
QY 626 ValProValGlyProGlnAaenPheSerLeuAlaIle 637
DB 2676 CTGCGCGGCGGCGCTGAAGAGCTCGCATTTCTCGTC 2711

```

```

RESULT 4
US-11-156-062-22
/ Sequence 22, Application US/1156062
/ Publication No. US20050281773A1
/ GENERAL INFORMATION:
/ APPLICANT: Wieland, Susanne
/ APPLICANT: Polanyi-Bald, Laura
/ APPLICANT: Prueßer, Inken
/ APPLICANT: Stehr, Regina
/ APPLICANT: Maurer, Karl-Heinz
/ TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
/ FILE REFERENCE: HEK-0134 / H5698
/ CURRENT APPLICATION NUMBER: US/11/156, 062
/ PRIOR FILING DATE: 2005-06-17
/ PRIOR APPLICATION NUMBER: PCT/EP2003/014127
/ PRIOR FILING DATE: 2003-12-20
/ PRIOR APPLICATION NUMBER: DE 102 60 903.9
/ PRIOR FILING DATE: 2002-12-20
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 22
/ LENGTH: 1140
/ TYPE: DNA
/ ORGANISM: Bacillus licheniformis
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1137)
/ NAME/KEY: mat_peptide
/ LOCATION: (316)..(1140)
US-11-156-062-22

```

```

Alignment Scores:
Pred. No.: 9,52e-17 Length: 1140
Score: 260.50 Matches: 120
Percent Similarity: 36.8% Conservative: 62

```

```

Best Local Similarity: 24.3% Mismatches: 157
Query Match: 8.0% Indels: 155
DB: 14 Gaps: 19
US-10-784-870-4 (1-639) x US-11-156-062-22 (1-1140)
QY 1 MetArgLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrVal 20
DB 4 ATGAGGAAAGAGTTTGGCTTGGATGTCGACCTTCATCGCTCGTTCCAGATG 63
QY 21 AlaLeuAaen-----ProSerAlaGlyAspAlaArgThrPheApbLeuApbPheLys 38
DB 64 GCATTCAGCATTCGCCCTTCTGCTGCTCAACCGCGCAAAATGTTGAAAGATAT-- 120
QY 39 GlyIleGlnThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAla 58
DB 121 -----ATTGCGGATTT-----AAGTCAGAGTCGAA 147
QY 59 AlaPheLeuValGluSerGluAenValLysLeuLeuLysGlyLeuLeuLysLysLeu 78
DB 148 ACCGATCTGTCAAAAGACATCATCAAGAGCGCGGAAAGTGCAGG----- 201
QY 79 ThrValProAlaAaenLysLeuHsileValGlnPheAaenGlyProIleLeuGln 98
DB 202 -----CAATTGGA-----ATCATCAACGC 222
QY 99 ThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAaenPyrIleProAspTyrAla 118
DB 223 GCAAAAGCGAGCTTACAAAGAGAGCGCTTAAGGAATCAAAATGATCCGATGTCGT 282
QY 119 TyrIleValGluTyrGluGluAaenValGlnSerLysValArgSerIleGluHsileValGlu 138
DB 283 TATGTGAAGAG-----GATCATGTG-- 303
QY 139 SerValGluProTyrLeuProLysTyrLysIleAaenProGlnLeuPheThrLysGlyAla 158
DB 303 ----- 303
QY 159 SerThrLeuValLysAlaLeuAlaLeuAaenThrLysGlnAaenLysGluValGlnLeu 178
DB 303 ----- 303
QY 179 ArgGlyIleGluGluIleAlaGlnTyrValAlaSerAaenAaenHsileTyrIleThrAla 198
DB 304 -----GCCCATGCGCTAGCGCA 321
QY 199 LysProGluTyrLysValMetAaenAaenValAlaArgGlyIleValLysAlaAaenValAla 218
DB 322 ACCGTTCTTACGCGCAT-----CTCTCATTAAGCGCAAAAGTG 363
QY 219 GlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAaenThrGlyLeu 238
DB 364 CAGGCTCAA--GGCTTAAGGAGCGAATGTAAAGTAGCCGCTGATACAGAAATC 420
QY 239 AspThrGlyArgAaenAaenSerSerMetHsileValPheArgGlyLysIleThrAlaLeu 258
DB 421 CAAGCTTCATCCGAGACTTGAACGTAGTCGCGCA-----GCAAGC 462
QY 259 TyrAlaLeuGlyArgThrAaenAaenAlaAaenAaenThrAaenGlyHsileGlyThrHsileValAla 278
DB 463 TTGTGGCTGGCGAAGCTTATAC--ACCGAGCGCAACGACACGCGCACTGTGCG 519
QY 279 GlySerVal-----LeuGlyAaenGlyAlaThrAaenLysGlyMetAlaProGlnAlaAaen 296
DB 520 GGTACAGTACGTGCGCTGACAAATCAAGGGGTATATAGGGGTGGCGCAAGCGATATCC 579
QY 297 LeuValPheGlnSerIleMetAaenAaenSerGlyGlyLeuGlyGlyLeuProSerAaenLeu 316
DB 580 TTGTACGCGTTAAAGTACTGAATTCAGCGCA-----AGCGATCATACAGCGCGCATTT 633
QY 317 GlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHsileThrAaenSerTyrGly 336
DB 634 GTACGCGAATGAGTGGCGAGCAACAAAGCGATGATATATGAGCTTGGG 693

```

```

Qy 337 AlaAlaValAsnGlyValAlaThrThrThrAspSerArgAsnValAspAspTyrValArgLys 356
    |||
Db 694 GGAGCA-----TCAGGCTGACAGCGATGAAACAGGCACTCGACATGCAATGATGCA 744
Qy 357 AsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArgProAsnGly-----GlyThr 374
    |||
Db 745 AGAGGGGTGTGCTGTAGCTGCAGCAGGGAAACAGGGATCTTCAGAGAAACGAAATACA 804
Qy 375 IleSerAlaProGlyThrAlaValAsnAlaIleThrValGlyAlaThrGluAsnLeuArg 394
    |||
Db 805 ATAGGCTATCTCGCGAAATACGATTTGTCATCGCTGTGGTGGCG-----849
Qy 395 ProSerPheGlySerTyrAlaAspAsnIleAsnIleValAlaGlnPheSerSerArgGly 414
    |||
Db 850 -----GTAAGCTTAAACAGACAGACAGCTTCATTTTCAGCGCTGGA 891
Qy 415 ProThrLysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSer 434
    |||
Db 892 -----GCAGAGCTTGAAGTCATGCTCTCGGCGCAGCGGATATACAGC 933
Qy 435 AlaArgSerSerLeuAlaProAspSerSerPheTPrAlaAsnHisAspSerLysTyrAla 454
    |||
Db 934 ACTTACCCACG-----AACACTTATGCA 957
Qy 455 TyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAla-----471
    |||
Db 958 ACATTGAACGAAACGTCATGCTTCTCTCATGTCAGGAGACAGCAGCTTGTGATCTTG 1017
Qy 472 -----GlnLeuArgGluHisPheValLysAsnArg 481
Db 1018 TCMAAACATCCGAACTTCAGCTTCACAAATCCGGAACGCT 1059

```

RESULT 5
US-11-020-602-1

Sequence 1, Application US/11020602
Publication No. US20060024764A1

GENERAL INFORMATION:

APPLICANT: Batell, David

APPLICANT: Harding, Fiona

TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND

TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME

FILE REFERENCE: GCS27C2

CURRENT APPLICATION NUMBER: US/11/020,602

PRIOR FILING DATE: 2004-12-22

PRIOR APPLICATION NUMBER: US 09/500,135

PRIOR FILING DATE: 2000-02-08

PRIOR APPLICATION NUMBER: US 09/060,872

PRIOR FILING DATE: 1998-04-15

NUMBER OF SEQ ID NOS: 240

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 1

LENGTH: 1495

TYPE: DNA

ORGANISM: Bacillus amyloliquefaciens

FEATURE:

NAME/KEY: mat peptide

LOCATION: (417)..(1495)

FEATURE:

NAME/KEY: CDS

LOCATION: (96)..(1244)

FEATURE:

NAME/KEY: misc feature

LOCATION: (582)..(584)

OTHER INFORMATION: The num at positions 582 through 584 which in a

OTHER INFORMATION: preferred embodiment (aac) is to code for

OTHER INFORMATION: asparagine, but which may also code for proline.

FEATURE:

NAME/KEY: misc feature

LOCATION: (585)..(587)

OTHER INFORMATION: The num at positions 585 through 587 which in a

OTHER INFORMATION: preferred embodiment (ccc) is to code for proline,

OTHER INFORMATION: but which may also code for asparagine.

```

FEATURE:
NAME/KEY: misc feature
LOCATION: (597)..(599)
OTHER INFORMATION: The num at positions 597 to 599 which in a
OTHER INFORMATION: preferred embodiment (aac) is to code for
OTHER INFORMATION: asparagine, but which may also code for aspartic acid.
NAME/KEY: misc feature
LOCATION: (678)..(680)
OTHER INFORMATION: The num at positions 678 through 680 which in a
OTHER INFORMATION: preferred embodiment (gca) is to code for
OTHER INFORMATION: alanine, but which may also code for serine.
FEATURE:
NAME/KEY: misc feature
LOCATION: (681)..(683)
OTHER INFORMATION: The num at positions 681 through 683 which in a
OTHER INFORMATION: preferred embodiment (tca) is to code for serine,
OTHER INFORMATION: but which may also code for alanine.
FEATURE:
NAME/KEY: misc feature
LOCATION: (708)..(710)
OTHER INFORMATION: The num at positions 708 through 710 which in a
OTHER INFORMATION: preferred embodiment (gct) is to code for
OTHER INFORMATION: alanine, but which may also code for aspartic acid.
FEATURE:
NAME/KEY: misc feature
LOCATION: (711)..(713)
OTHER INFORMATION: The num at positions 711 through 713 which in a
OTHER INFORMATION: preferred embodiment (gac) is to code for
OTHER INFORMATION: aspartic acid, but which may also code for alanine.
FEATURE:
NAME/KEY: misc feature
LOCATION: (888)..(890)
OTHER INFORMATION: The num at positions 888 through 890 which in a
OTHER INFORMATION: preferred embodiment (acc) is to code for
OTHER INFORMATION: threonine, but which may also code for serine.
FEATURE:
NAME/KEY: misc feature
LOCATION: (891)..(893)
OTHER INFORMATION: The num at positions 891 through 893 which in a
OTHER INFORMATION: preferred embodiment (tcc) is to code for
OTHER INFORMATION: serine, but which may also code for threonine.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1167)..(1169)
OTHER INFORMATION: The num at positions 1167 through 1169 which in
OTHER INFORMATION: a preferred embodiment (gaa) is to code for
OTHER INFORMATION: glutamic acid, but which may also code for glutamine.
US-11-020-602-1

```

Alignment Scores:

Pred. No.:	5,27e-15	Length:	1495
Score:	245.00	Matches:	116
Percent Similarity:	34.2%	Conservative:	53
Best Local Similarity:	23.5%	Mismatches:	147
Query Match:	7.5%	Indels:	178
DB:	14	Gaps:	18

US-10-784-870-4 (1-639) x US-11-020-602-1 (1-1495)

```

Qy 1 MetArgLysLysValPheLeuSerValLeuSerAlaAlaIleLeuSerThrVal 20
    |||
Db 96 ATGAGAGCAAAAAGATGATGATGCTGTTCTTTCAGGCTTAACTTTCATGATG 155
Qy 21 AlaLeuAsnAsnProSerAlaGlyAspAlaArgThrPheAspLeuAspPheLysGlyIle 40
    |||
Db 156 GCGTTCGCGACATCTCTGCCAGGCG-----185
Qy 41 GlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGlyAlaAlaAlaPhe 60
    |||
Db 186 -----GCAGGGGAATCAACCGGGAAG-----209
Qy 61 LeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLeuLysLysLeuGluThrVal 80

```



```

Db      210 -----AAATATATTGCGGTTTAAACAGCAAGATGAGACGATG 248
Qy      81  ProAlaAsnAsnLysLeuHisIleValGlnPheAsnGlyProIleLeuGlnGluThrLys 100
Db      249  AGCGCGCGTAAAGAAAGAAAGATGCAATTTCTGAAAAGAGC----- 287
Qy      101  GlnLysLeuGlnIuThrThrGlyAlaLysIleLeuAspTyrIleProAspTyrAlaTyrIle 120
Db      287 ----- 287
Qy      121  ValGlnTyrGlnGlyAspValGlnSerLysValArgSerIleGlnHisValGlnSerVal 140
Db      288 -----GGAAAGTGCAGAAAGCAATCAAAATATGTAGC----- 320
Qy      141  GluProTyrLeuProLysTyrLysIleAspProGlnLeuPheThrLysGlyAlaSerThr 160
Db      321 -----GCAGCTTCAGCT 332
Qy      161  LeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLysGlnValGlnLeuAspGly 180
Db      333  ACATTAAACGAAAAAGCTGTA-----AAAGAAATGAAAAAGAACCCGAGCGTCTTAC 386
Qy      181  IleGlnGlnIleAlaGlnTyrValAlaSerAsnAspValHisTyrIleThrAlaLysPro 200
Db      387  GTTGAAGAA-----GATCAGTAGCA-----CATCGCTACGGCAGTCCGCTG 428
Qy      201  GluTyrLysValMetAsnAspValAlaArgGlyIleValLysAlaAspValAlaGlnSer 220
Db      429  CCTTACGGCGTATCACAA-----ATTAAAGCCCTT---GCTTCGAC 467
Qy      221  SerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAlaAspThrGlyLeuAspThr 240
Db      468  TCTCAAGGCTACACGTGATCAAAATGTTAAAGTAGCGTTATTCACAGCGGTATCGAT--- 524
Qy      241  GlyArgAsnAspSerSerMetHisGlnAlaPheArgGlyLysIleThrAlaLeuTyrAla 260
Db      525 -----TCTCTCATCTCTGATTAAAGTAGACGAGCGGACGACAGATGTT 569
Qy      261  LeuGlyArgThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySer 280
Db      570  CCTTCTGAACANNNNNTTCCAAACNNNNACTCTCAGGAACTCAGCTTGCGGCAC 629
Qy      281  Val-----LeuGlyAsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuVal 298
Db      630  GTTGCAGGCTTTAATACCTCAATCGGTATTAAGGGGTGCGCAAGCANNNNCTTTAC 669
Qy      299  PheGlnSerIleMetAspSerSerGlyGlyLeuGlyGlyLeuProSerAsnLeuGlnThr 318
Db      690  GCTGTAAAGTTCTCGTNNNNNGGT----- 716
Qy      319  LeuPheSerGlnAlaPheSerAlaGlyAlaArgIleHisThrAsnSerTyrGlyAlaAla 338
Db      717 -----TCCGGCCAAATACAGCTGGATC 737
Qy      339  ValAsnGlyAlaLeuTyrThrAspSerArgAsnValAsp----- 351
Db      738  ATTAAACGAATCGAGTGGCGGATCGCAACAAATATGACGTTATTAACATGAGCCTCGGC 797
Qy      352 -----AspTyrValArgLysAsnAspMet 359
Db      798  GGAACCTTCGTTCTGCTGCTTTAAAGCGGAGCTTGATTAAGCCGTTGATCCGCGCTC 857
Qy      360  ThrIleLeuPheAlaAlaGlyAsnGlyLysArgProAsnGly-----GlyThrIleSerAla 377
Db      858  GTAAGCTGGTGGCAGCCGCTAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 917
Qy      378  ProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGlnAsnLeuArgProSerPhe 397
Db      918  CCTGTAAATACCTTCTGCTGATGAGTGAAGGCGCT----- 953
Qy      398  GlySerTyrAlaAspAsnIleAsnHisIleValAlaGlnPheSerSerArgGlyProThrLys 417

```

```

Db      954 -----GTTGACAGCAGCAACAAAGACATCTTTCTCAAGCGTAGACCT----- 998
Qy      418  AspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArgSer 437
Db      999 -----GAGCTTGATGTCATGCGACCTGGC-----GTATTCATCCAAAGC 1037
Qy      438  SerLeuAlaProAspSerSerPheThrAlaAsnHisAspSerLysTyrAlaTyrMetGly 457
Db      1038  ACCCTTCT-----GAAACAAATACGGGGCGTACAC 1070
Qy      458  GlyThrSerMetAlaThrProIleValAlaGlyAsnValAla 471
Db      1071  GGTACGTCAATGCAATCTCCGACGTTGCCGAGCGGCGCT 1112

RESULT 6
US-11-136-527-605
; Sequence 605, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Nouns, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 605
; LENGTH: 4216
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-605

Alignment Scores:
Pred. No.: 2,356-14 Length: 4216
Score: 244.00 Matches: 138
Percent Similarity: 38.3% Conservative: 87
Best Local Similarity: 23.5% Mismatches: 233
Query Match: 7.5% Indels: 130
DB: 14 Gaps: 27

US-10-784-870-4 (1-639) x US-11-136-527-605 (1-4216)
Qy      87  HisIleValGlnPheAsnGlyProIleLeuGlnGluThrLysGlnLysLeuGlnThr 106
Db      583  TATATTGTGCTTCAACGGAATACCTCACAGCAAAAGTAAACATTTATTTCAGAT 642
Qy      107  GlyAlaLys-----IleLeu 111
Db      643  GCTTAAAGCAAGTGAAGTGAACAAGTGAATTAATCTCGAACAACCATCTAGT 702
Qy      112  AspTyrIleProAspTyrAlaTyrIleValGlnGlyLysAspValGlnSerLysVal 131
Db      703  GACTACCTTAAGTATTTAG-----GTGATTCAGATTAAGAAAGAGAGAGAGGCGGCTG 759
Qy      132  ArgSerIleGlnHisIleValGlnSerValGluProTyrLeuProLysTyrLysIleAspPro 151
Db      760  CTCACACTGAAGATCACCAACCAACATCAAGCGGTGACACCCACGGAAGATCTTCGT 819
Qy      152  GlnLeuPheThrLysGlyAlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGln 171
Db      820  TCCCTGAAGTTGCTGAATCCGACCCCATGTGCCC-----TGTATGAGACCCCGGTG 873
Qy      172  AsnAsnLysGlnValGlnLeuArgGlyIleGlnGlnIleAlaGlnThrValAlaSerAsn 191
Db      874  AGCAGAGATGCGACATCACGTCCCTGAAAAGAGCCAGTCTCCCTGGGCTTGCA 933
Qy      192  AspValHisTyrIleThrAlaLysProGluTyrTyrLysValMetAsnAspValAlaArgGly 211
Db      934  TTCTGCAATGCAACAGAAAGCAATTCAGTCAACATTCGTAAGCCATTCTCGCCAG 993

```

```

QY 212 lle-----VallysaiaapValaIaGlnSerSerTrgIyLeuTrgIyGlnGly 228
   :::::|||||
Db 994 GTTGGCCAGACATTCGACGGAGATGCTTGG---CAGATGGAGATACACAGGCTCAAT 1050
QY 229 GlnIleValaIaValaIaapTrhTrgIyLeuapTrhTrgIyAraAapSerSerMetHis 248
   |||||
Db 1051 GTCAGGGTTGCTGTTTGTACTGCGGCTC-----AGTGAAGACAT 1092
QY 249 GluaIaPheAraGlyIySerIleThraIaLeuTrgIyAraIaLeuGlyAraThraAapAaAa 268
   |||||
Db 1093 CCACATTCACAGAAAT-----GTGAAGAAAGAAACCACTGGACCAAT 1134
QY 269 -----AapThraAengIyHiAegIyThraIaValaIaGlySerValaLeuGly 283
   |||||
Db 1135 GAGCGGACCCCTGGACATGGCGCTGGCCATGGACATGCTTGCAGCT---GTGATGGC 1191
QY 284 AaenGlyAraThraAaenGlyMetAaProGlnAaAaAaLeu---ValPheGlnSerIle 302
   :::::|||||
Db 1192 AGCATGAGAGAGTGCAGAGGATTTGGCCCGCATGACAGCTGCACATCTTCAGGGCTTT 1251
QY 303 MetAapSerSerGlyIyLeuGlyIyLeuProSerAaAaLeuGlnThraLeuPheSerGln 322
   :::::|||||
Db 1252 ACCAACATACAG-----GTGCTTACACGCTTGGCTTTTGGAT 1290
QY 323 AlaPheSerAaGlyAraIaAraGlyIeHisThraAaSerTrpGlyAraIaAaValaAengIyAa 342
   |||||
Db 1291 GCCTTCACTATGACCTCCAAAGAAAGATKAGACGTTCTGAACCTTACGATCGGCGGCT 1350
QY 343 TrpTrhThraAapSerAraAaAaValaAaAaPtyr-----ValAraGlyAaAapMetTrh 360
   |||||
Db 1351 GACTTCATGATCAACCCCTTTGTTCAGAGTATGGAAATTTAACAGCGAACATGTAATC 1410
QY 361 IleLeuPheAaIaAaIaGlyAaAengIyAraGProAaengIyIyThraIleSerAaIaProGlyTrh 380
   :::::|||||
Db 1411 ATGCTTTCTGCTATTTGGCAATGATGACCTCTCTATGGACCTCTGAATTAACCTTGCTGAT 1470
QY 381 AlaIyAaAaAaIaIeThraValaGlyAaThraGlnAaAaAraGProSerPheGlySerTr 400
   :::::|||||
Db 1471 CAGATGATGATGATTTGAGTGGGTGGCATTTGAC-----TTT 1506
QY 401 AlaAaAaAaIleAaAaIaValaIaGlnPheSerSerAraGly----- 414
   |||||
Db 1507 GAAGACAAACATC-----GCCCGCTTCTCTTCAGGGGAAATGATCACTGGGAACATA 1557
QY 415 ProThraIyAaPgyAraGlyIyLeuProAaPValaMetAaIaProGlyThraTrpIleLeuSer 434
   |||||
Db 1558 CCGGAGAGCTATGCTGCTGTAAGCTTACATTTGTC-----ACCTATGCTCTGGA 1608
QY 435 AlaAraGSerSerLeuAaIaProAaPSerSerPheTrpAaAaAaIaAaPSerTrpAraIa 454
   |||||
Db 1609 GTGCGGGGTTCTGCTGTGAAGGGGGCTGC-----GGT 1641
QY 455 TyrMetGlyGlyThraSerMetAaThraProIleValaIaGlyAaAaValaIaGlnLeuAraG 474
   :::::|||||
Db 1642 GCACCTCAGGAGACAGTGTCCGCTCCCAAGTGTGCTGGGGCTGATCACTTTGTAATA 1701
QY 475 GlnHisPheValaIyAaAaAraGlyIleThraProIyProSerLeuLeuAaIaAaLeu 494
   |||||
Db 1702 AGCACAGTAACAGAACGGGAGCTAGTG-----AATCTGCGCAAGTGTGAAGAACGCTTGG 1755
QY 495 IleAaGlyAaIaAaAaPValaGlyLeuGlyTrpProAaengIyAaAa-----GlnGly 511
   :::::|||||
Db 1756 ATAGCATCAGCCCGGAGACTT-----CCTGGGTCAACATGTTTGAGCAAGGC 1803
QY 512 TrpGlyAraGValaThraAaPlyAaSerIleAaAaVala-----AaTrpValaAaGlnSer 529
   |||||
Db 1804 CATGCAAGTGGATCTACTGAGACCTATGACATCTCACAGACTATTAACCGGAGCG 1863
QY 530 SerAaLeuSerThraSerGlnAaIaThraTrpThraAaIaThraAaGlyAaPtyr 549
   |||||
Db 1864 AGCCG-----AGTCTT 1875
QY 550 LeuIyAaIleSerLeuVala-----TrpSerAaPalaProAaIaSer 562

```

```

Db 1876 ACCTACATGACCTGATCGATGCTCCCTTACATGAGCCCTTGTCTCCAGCCATCTAC 1935
QY 563 ThraAaIaSerValaThraLeuValaAaAaPheAaAaPheAaValaIleThraAaProAaengIy 582
   |||||
Db 1936 TATGAGGAATGCCAACATGTTTAA-----GTCAACATCTCTCAATGCG 1980
QY 583 ThraGlyTrpValaGlyAaAaAaPheAaAaAaPheAaAaAaPheAaAaAaAaAaAa 598
   |||||
Db 1981 ATGGAGTTACAGAAAGATGTGTGATAGCT-----GAGTGGCGAACCTATTTA 2031
QY 599 ---AapGlyAraGAAaAaAaValaGlnAaAaAaPheIleAaAaSer-----ProGln 613
   :::::|||||
Db 2032 CCACGAATGAGACACATTTGAAGTGGCTTCTCTACTCTCTCACTGATGTTGGGCTTGG 2091
QY 614 SerGlyTrpTrpTrpIleGlnVal 621
   |||||
Db 2092 TCAGGTTACCTTGCATCTCCATT 2115
RESULT 7
US-11-122-329-113
/ Sequence 113, Application US/11122329
/ Publication No. US20060019272A1
/ GENERAL INFORMATION:
/ APPLICANT: Geraci, Mark
/ APPLICANT: Bull, Todd
/ APPLICANT: Voelkel, Norbert
/ APPLICANT: Coldren, Chris
/ TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
/ FILE REFERENCE: 2848-54
/ CURRENT APPLICATION NUMBER: US/11/122,329
/ PRIOR FILING DATE: 2005-05-03
/ PRIOR APPLICATION NUMBER: 60/568,129
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 113
/ LENGTH: 4338
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-122-329-113
Alignment Scores:
Pred. No.: 3,08e-14 Length: 4338
Score: 243.00 Matches: 140
Percent Similarity: 38.6% Conservative: 89
Best Local Similarity: 23.6% Mismatches: 224
Query Match: 7.4% Indels: 140
DB: 14 Gaps: 28
US-10-784-870-4 (1-639) x US-11-122-329-113 (1-4338)
QY 87 HisIleValaGlnPheAaengIyProIleLeuGlnGlnThraIySerGlnIyLeuGlnThra 106
   ::|||
Db 662 TATATTTGGCTTTCANTGATTAATTACAGCCAAACTAGAAATTCATTATTTCAGT 721
QY 107 GlYAlaIyS----- 111
Db 722 GCCCTGAAGAGCACTGAAGTGAACATTTGAGATTATTACTCGAAACAATTCATCCAGT 781
QY 112 AspTrpIleProAaPtyrAraIaTrpIleValaGlnTrpGlnIyAaPValaGlnSerIyVal 131
   |||||
Db 782 GACTTACCTTGTGTGATTTTGG---GTGATTCAGATTAAGAAAGAAAGAAAGCGGGGCTG 838
QY 132 ArgSerIleGlnHisValaGlnSerValaGlnProTrpTrpLeuProIyTrpIyAaIleAaPPro 151
   ::|||
Db 839 CTAACTTGAAGATATCAACAACATCAAGGCTACGCGCCCAAGAAAGTCTTTCGT 898
QY 152 GlnLeuPheThraIySerGlyAaSerThraLeuValaIyAaIaLeuAaAaPtyrIyGln 171
   |||||
Db 899 TCCCTC-----AAGTATGCTGAATGACCCCAAGTACCTCGATCAATGAAGAACCGGCTGG 952

```

```

QY 172 AsnAsnLysGluValGlnLeuArgLysIleGluGlnIleAlaGlnIleValAlaSerAsn 191
   |||
   |||
   |||
Db 953 AGCAGAGAGTGGCATATCATCGTCCCTCGGAGAGACCCCTCTCCCTGGCTTGGC 1012
QY 192 AspValHisThrIleThrAlaLysProGlnIleValMetAsnAspValAlaArgGly 211
   |||
   |||
   |||
Db 1013 TTCTGCGATCGTACGGGAGGAGCATTCGAGCAGACGCGCTGCGAGAGCCGCGCAG 1072
QY 212 Ile-----VallysaAspValAlaGlnSerSerIleValLeuIleGlnGly 228
   |||
   |||
   |||
Db 1073 GTTCCCGCAGACCTGCGAGCATGTCCTCG--CAGATGGATATACAGGCTAT 1129
QY 229 GlnIleValAlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetHis 248
   |||
   |||
   |||
Db 1130 GTTAGAGTCTGCTTTTGGACCTGGGCTG-----AGCGAGAGCAT 1171
QY 249 GluAlaPheArgLysIleThrAlaLeuTyraLeuGlyArgThrAsnAsnAlaAsn 268
   |||
   |||
   |||
Db 1172 CCCCATTCATAAAT-----GTGAGAGAGAGAACCACTGGAGCCAC 1213
QY 269 -----AspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGly 283
   |||
   |||
   |||
Db 1214 GACGAGACCTGAGCATGGGTTGGGCTGAGCATGTCATTCGTGGCAGGT--GTGATGGC 1270
QY 284 AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeu--ValPheGlnSerIle 302
   |||
   |||
   |||
Db 1271 AGCATAGGAGGAGTCCAGAGATTGCTCCAGATGCGAATTCATTCACAGGCTCTT 1330
QY 303 MetAspSerSerGlyLeuGlyLeuProSerAsnLeuGlnThrLeuPheSerGln 322
   |||
   |||
   |||
Db 1331 ACCAATATATCAG-----GATCTTACACATTCGTGTTTGGAC 1369
QY 323 AlaPheSerAlaGlyAlaGlyIleHisThrAsnSerTrpGlyAlaAlaValAlaGlyAla 342
   |||
   |||
   |||
Db 1370 GCCTTCACATATCCATTTTAAAGAGATGACGTGTTAACTCAGCATCGCGGCGC 1429
QY 343 TyrTrpThrAspSerArgAsnValAspAspTyr-----ValArgLysAsnAspMetThr 360
   |||
   |||
   |||
Db 1430 GACTTCATGATCATCGCTTGTGTCAGAGTGGAGATTCACAGTTCACATATCATC 1489
QY 361 IleLeuPheAlaAlaGlyAsnGluArgProAsnGlyGlyThrIleSerAlaProGlyThr 380
   |||
   |||
   |||
Db 1490 ATGCTTTCGCTATGGCAGATGACGACCTCTTATGGCAGCTCGAATACCTGCTGAT 1549
QY 381 AlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSerPheGlySerTyr 400
   |||
   |||
   |||
Db 1550 CAATGATGCTGATTTGAGATGGCGCGCATTTGAC-----TTT 1585
QY 401 AlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGly----- 414
   |||
   |||
   |||
Db 1586 GAGATATATCATC-----GCCCGCTTTCTTCAAGGGAGATGACTACCTGGAGCTA 1636
QY 415 ProThrLysAspGlyArgIleLysProAspValMetAlaProGlyThrTrpIleLeuSer 434
   |||
   |||
   |||
Db 1637 CCAAGAGGCTACGCTGCGATGAGAACCTGACATGTC-----ACCTATGGTGTGCG 1687
QY 435 AlaArgSerSerLeuAlaProAspSerSerPheTrpAlaAsnHisAspSerLysTyraAla 454
   |||
   |||
   |||
Db 1688 GTCCGGGGTTCGCTCGTGAAGGGGGTGC-----CGG 1720
QY 455 TyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArg 474
   |||
   |||
   |||
Db 1721 GCCCTCTCAGGGACGAGTGTCTCTCCAGTGGTGGAGTCTGTCACCTTGTTAGTG 1780
QY 475 GluHisPheValLysAsnArgGlyIleThrProLysProSerLeuLeuLysAlaAlaLeu 494
   |||
   |||
   |||
Db 1781 AGCAGAGTCCAGAGAGGTGAGCTGCTG-----ATTCGCCGACGATGAGAGAGCCCTG 1834
QY 495 IleAlaGlyAlaAlaAspValGlyLeuGlyTyrProAsnGlyAsn-----GlnGly 511
   |||
   |||
   |||
Db 1835 ATCGCTCAGCGCCGAGGCTC-----CCCGGGGTCAACATGTTTGAGCAAGCG 1882
QY 512 TrpGlyArgValThrLeu-----AspLys 519

```

```

Db 1883 CAGCGCAGCTCATCTGCTCAGAGCCTATCAGATCTCAACAGCTACAGCCACAGCA 1942
QY 520 SerLeuAsnValAlaTyraValAsnGlySerSerAla-----IleSerThrSer 535
   |||
   |||
   |||
Db 1943 AGTTTGAAGCCCGACGATCATATGATGAGTGTCCCTACATGTCGCCCTACTGCTCC 2002
QY 536 GlnLysAlaThrTrpThrPheThrAlaThrAlaGlyLysProLeuLysIleSerLeuVal 555
   |||
   |||
   |||
Db 2003 CAGCCCATCTACAT-----GAGAGATGCG----- 2029
QY 556 TrpSerAspAlaProAlaSerThrThrAlaSerValThrLeuValAsnAspLeuAspLeu 575
   |||
   |||
   |||
Db 2030 -----ACAGTGTATATGTCACATCTCCACAGCGGATGGAGTCTC 2068
QY 576 -----ValIleThrAlaProAsnGlyThrArgTyraGlyLysAsnAspPheSerAla 592
   |||
   |||
   |||
Db 2069 ACAGAGAAATGTAGATTAAGCTGAGCTGGCAGCCCTATTTG----- 2110
QY 593 ProPheAspAsnAsnTrpAspGlyArgAsnAsnValGluAsnValPheIleAsnSer-- 611
   |||
   |||
   |||
Db 2111 -----CCACAGAACGAGAGC-----ACATGGAAGTTCCTTCTCTACTCTCG 2155
QY 612 -----ProGlnSerGlyThrTrpThrIleGluVal 621
   |||
   |||
   |||
Db 2156 GTCTTATGCTTGTGTCGGGCTACCTGGCCATCTCCAT 2194

RESULT 8
US-11-047-380-4
; Sequence 4, Application US/11047380
; Publication No. US20060068406A1
; GENERAL INFORMATION:
; APPLICANT: AEPHOLTER, JOSEPH A.
; APPLICANT: COX, ANTHONY
; APPLICANT: NESS, JON E.
; APPLICANT: CARR, BRIAN
; TITLE OF INVENTION: SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED
; TITLE OF INVENTION: RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION
; FILE REFERENCE: 0165,551US
; CURRENT APPLICATION NUMBER: US/11/047,380
; PRIOR FILING DATE: 2005-01-31
; PRIOR APPLICATION NUMBER: 09/721,507
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 09/656,549
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/185,244
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,815
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/186,247
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/186,482
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Subtilisin B
; NAME/KEY: modified base
; LOCATION: (1)..(16)
; OTHER INFORMATION: a, c, t, g, other or unknown
; NAME/KEY: modified base
; LOCATION: (1163)..(1180)
; OTHER INFORMATION: a, c, t, g, other or unknown
US-11-047-380-4
Alignment Scores:

```

Pred. No.:	7,18e-15	Length:	1160
Score:	242.50	Matches:	123
Percent Similarity:	37.3%	Conservative:	55
Best Local Similarity:	25.8%	Mismatches:	154
Query Match:	7.4%	Indels:	145
DB:	11	Gaps:	19

US-10-784-870-4 (1-639) X US-11-047-380-4 (1-1180)

Qy	1	MetArglyVal	Val	Phe	Leu	Ser	Val	Leu	Ser	Ala	Ala	Leu	Ser	Thr	Val	20								
Db	17	GTGAGAA	GCAAAAA	TTGTG	GATCAG	GTGTT	GTGTT	GTGTT	GTGTT	GTGTT	GTGTT	GTGTT	GTGTT	GTGTT	GTGTT	76								
Qy	21	Ala	Leu	Asn	Asn	Pro	Ser	Ala	Arg	Val	Arg	Thr	Phe	Asp	Leu	Asp	Phe	Val	Leu	Val	40			
Db	77	GCGTT	CAGCA	AACTGT	CTCG	CAGCG	CGTGG	CGCGG	CGGAAAA	AGCA	GTCAC	AGAA	AGAA	ATTCAT	TTT	13								
Qy	41	Gln	Thr	Thr	Thr	Asp	Val	Ser	Val	Phe	Ser	Val	Gln	Arg	Gln	Thr	Thr	Val	Ala	Ala	Phe	60		
Db	137	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	16		
Qy	61	Leu	Val	Glu	Ser	Glu	Asn	Val	Val	Ser	Leu	Leu	Val	Gly	-----	Leu	Leu	Val	Ser	Glu	Val	78		
Db	170	TCCGC	CAAGAAA	AAAGAT	GTATT	ATTC	GAAAA	AGCGG	CAAA	AGCA	AGTTCA	AAAGCA	ATTT	TAAG	22									
Qy	79	Thr	Val	Pro	Ala	Asn	Asn	Val	Ser	Val	Ser	Val	Ser	Val	Arg	Ser	Val	Arg	Ser	Val	Arg	98		
Db	230	TATGT	TAA	CCCG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	24		
Qy	99	Thr	Val	Arg	Gln	Val	Ser	Glu	Thr	Thr	Thr	Val	Ala	Val	Leu	Asp	Thr	Leu	Pro	Asp	Thr	Val	11	
Db	242	GCCGC	CAACCA	ATTG	ATG	AAAA	AGCTG	TAA	AAAGAT	TAA	AAAA	AGATTC	CGAC	CGTT	CA	30								
Qy	119	Tyr	Val	Glu	Val	Arg	Glu	Val	Arg	Ser	Val	Arg	Ser	Val	Arg	Ser	Val	Arg	Ser	Val	Arg	136		
Db	302	TATGT	GAAGA	AAAT	CAT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	34		
Qy	139	Ser	Val	Glu	Pro	Tyr	Leu	Pro	Tyr	Val	Ser	Val	Ser	Val	Ser	Val	Ser	Val	Ser	Val	Ser	Val	156	
Db	341	TCTGT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	35		
Qy	159	Ser	Thr	Leu	Val	Val	Ala	Leu	Ala	Leu	Asp	Thr	Val	Gln	Asn	Asn	Val	Glu	Val	Gln	Leu	176		
Db	352	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	35		
Qy	179	Arg	Gly	Val	Glu	Glu	Val	Leu	Ala	Gln	Tyr	Val	Ala	Ser	Asn	Asp	Val	Ile	Tyr	Val	Leu	Thr	Ala	196
Db	353	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	36	
Qy	199	Lys	Pro	Glu	Tyr	Val	Val	Met	Asn	Asp	Val	Ala	Arg	Gly	Val	Leu	Val	Val	Asp	Val	Ala	216		
Db	368	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	37		
Qy	219	Gln	Ser	Ser	Tyr	Gly	Val	Tyr	Gly	Gln	Gly	Gln	Val	Ala	Val	Ala	Asp	Thr	Gly	Val	236			
Db	380	CTTCA	CTCTCA	AGGCT	CAACAG	CGCTCT	TAA	CGTAA	AAAGT	AGCTG	TAT	GCAC	AGCG	GAAT	43									
Qy	239	Asp	Thr	Gly	Val	Arg	Asn	Asp	Ser	Ser	Met	Ala	Gln	Ala	Phe	Arg	Gly	Val	Val	Leu	256			
Db	440	GACTC	TTTCAT	CTCG	ACTTAA	ACGTC	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	48			
Qy	259	Tyr	Val	Ala	Glu	Val	Arg	Thr	Asn	Asn	Ala	Asn	Asp	Thr	Asn	Gly	Val	Val	Val	Ala	278			
Db	482	TTCGT	ACTTTC	GAA	CAAA	CCCAT	ACCA	CCGAC	CGGCA	AGTCT	CA	CGGTA	CCAT	GTAG	CC	54								
Qy	279	Gly																						

```

Oy      317 GlnThrLeuPheSerClnIaPheSerAlaGlyAlaArgIleIleThrIleHisSerTyrGly 336
Db      656 ATTAACGGCATTTGAGTGGGCCATTTCACCAATATGATGATTTCACACATGACCTTGGC 715
Oy      337 AlaAlaValaAsnGlyAlaTyrThrThrIlePheSerArgAsnValaAspArgTyrValaArglys 356
Db      716 GGACCTACTGCT-----TCTACAGCGCTGAAACACGTCGTTGACAAAGCGTTTCC 766
Oy      357 AsnAspMetThrIleLeuPheAlaIaGlyAsnGluArgProAsnGly-----GlyThr 374
Db      767 ACGGATATGTCGTGGTCGCCGACGCGGAAACGAAGGTTCAATCCGAAAGCACACACACA 826
Oy      375 IleSerAlaProGlyThrAlaIysAsnAlaIleThrValaGlyAlaThrGluIleuAsnLeuArg 394
Db      827 GTCGGCTACCCCTGCAAAAATATCCTCTTACTATTGCAAGTGGTGGC----- 871
Oy      395 ProSerPheGlySerTyrAlaAspAsnIleAsnIleValaIleGlnPheSerSerArgGly 414
Db      872 -----GTAAACGACGACCAACAAAGACTTCATTCTCCACGGCGAGGT 913
Oy      415 ProThrIysAspGlyArgIleLysProAspValaIleMetAlaProGlyThrTyrIleLeuSer 434
Db      914 TCT-----GAGCTTATGTGATGGCTCTCCGGC-----GTGTCC 946
Oy      435 AlaArgSerSerLeuAlaProAspSerSerPheThrAlaAsnIleAspSerLysTyrAla 454
Db      947 ATCCAAACACACATTCCT-----GAGGACACTTACGGC 979
Oy      455 TyrMetGlyGlyThrSerMetAlaThrProIleValaIaGlyAsnValaAla 471
Db      980 GCTTATACGAACTCCATCGACCTCTTACAGTGTGGCGGACGACGACG 1030

```

RESULT 9
US-11-12

```

? Sequence 679, Application US/11128061
? Publication No. US20060003958A1
? GENERAL INFORMATION:
? APPLICANT: Melville, Mark W.
? APPLICANT: Charlebois, Timothy S.
? APPLICANT: Mounts, William M.
? APPLICANT: Hann, Louane E.
? APPLICANT: Sinacore, Martin S.
? APPLICANT: Leonard, Mark W.
? APPLICANT: Brown, Eugene L.
? APPLICANT: Miller, Christopher P.
? TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
? FILE REFERENCE: 01997.027701
? CURRENT APPLICATION NUMBER: US/11/128,061
? CURRENT FILING DATE: 2005-05-11
? PRIOR APPLICATION NUMBER: US 60/570,425
? PRIOR FILING DATE: 2004-05-11
? NUMBER OF SEQ ID NOS: 7285
? SOFTWARE: PatentIn version 3.3
? SEQ ID NO 679
? LENGTH: 4199
? TYPE: DNA
? ORGANISM: Cricetus gryseus
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (130)..(150)
? OTHER INFORMATION: n is a, c, g, or t
? FEATURES:
? NAME/KEY: misc feature
? LOCATION: (4177)..(4199)
? OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-679

Alignment Scores:
Pred. No.: 3,76e-14 Length: 4199
Score: 242.00 Matches: 138
Percent Similarity: 36.1% Conserves: 86
Dist Local Similarity: 23.5% Mismatches: 234

```

Query Match: 7.4% Indels: 130
DB: 14 Gaps: 27
US-10-784-870-4 (1-639) x US-11-128-061-679 (1-4199)

QY 87 HisIleValGlnPheAenGlyProIleLeuGlnGluThrLysGlnLysLeuGlnThr 106
DB 552 TATATTGTGGCTTCAACGATATTCTCACAGCCAAAGCTAAGAACTATTATTACAGT 611
QY 107 GYAlaLys-----IleLeu 111
DB 612 GCTCTGAAAGCAGTGAAGTAGACACTGGAATTATCTCGGAACAACCTCCAGT 671
QY 112 AAPTyrIleProApyrAlaTyrIleValGluTyrGlnLysApyValGlnInsLysVal 131
DB 672 GACTTACCTTGATGATTTTGAG--GTGATTCAGATTAAAGAAAGCAAGAAAGCCGGCTG 728
QY 132 ArgSerIleGlnHisValGlnSerValGluProTyrLeuProLysTyrLysIleAapPro 151
DB 729 CTCACACTTGAGATCATCAAAACATCAAGCGGGTGACCTCAAGCCAAAGCTTTCCGT 788
QY 152 GlnLeuPheThrLysGlnLysAserThrLeuValLysAlaLeuAlaLeuApyThrLysGln 171
DB 789 TCCTTGAAAGTTTCTGATCTGACCCCATGTGCGCA-----TGTAATGAACCTCGGCTG 842
QY 172 AapApyLysGlnValGlnLeuArgGlyIleGlnGlnIleAlaGlnIleValAlaAserAap 191
DB 843 AGCCAAAGTGGCAGTCATCAACACCCCTGAGAAAGCCGCTCTCCCTGGGCTCTGGA 902
QY 192 AapValHisTyrIleThrAlaLysProGluTyrLysValMetAapApyValAlaArgGly 211
DB 903 TTCTGCGATGCAACAGGAAGACATTCAGCCGGCGATGCTGAGAGCCATTCCTCGACAG 962
QY 212 Ile-----ValLysAlaApyValAlaGlnSerSerTyrGlnLysGlnGly 228
DB 963 GTTGCCAGACATTCAGGAGGAGATGTCGTGTG---CAGATGGGAATACAGAGTGTCAAT 1019
QY 229 GlnIleValAlaValAlaApyThrGlyLeuApyThrGlyArgAapApySerMetHis 248
DB 1020 GTCCAGGTTGCTGTGTTTGTACTGGGCTC-----AGTGAAAGCAAT 1061
QY 249 GlnAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGlyLysArgHisAapApySer 268
DB 1062 CCACACTTCAGAAAT-----GTGAGAGAGAAACCACTGACCAAT 1103
QY 269 -----AapThrAenGlyLysIleGlyThiValAlaGlySerValLeuGly 283
DB 1104 GAGCGGACCTGATGATGAGGCTGGCCATGCGACATTTGTCGACAGT---GTGATTCGC 1160
QY 284 AenGlyAlaThrAenLysGlyMetAlaProGlnAlaAenLeu---ValPheGlnSerIle 302
DB 1161 AGCATGAGGAGTCCAGGAGATTTGGCCAGATGAGCTGCACATCTTCCGGGCTCTTT 1220
QY 303 MetAapSerSerGlyLysLeuGlyLysLeuProSerAenLeuGlnThrLeuPheSerGln 322
DB 1221 ACCCAAAATCAG-----GTGCTTACACATCTGTGTTTGGAC 1259
QY 323 AlaPheSerAlaGlyAlaArgIleHisThrAenSerTyrGlyAlaAlaValAenGlyAla 342
DB 1260 GCTTTCACCTAGCTCCCTAAAGAAAGATTGATCTTAAACCTTAGCATCGGCGGCTT 1319
QY 343 TyrThrThrAapSerArgAenValAapApyr-----ValArgLysAenAapMetThr 360
DB 1320 GACTTCATGAGATCATCCCTTTGTTGCAAGGTGTGGAAATTAACACTTAACATGTAAATC 1379
QY 361 IleLeuPheAlaAlaGlyAenGlnLysProAenGlyLysThrIleSerAlaProGlyLys 380
DB 1380 ATGGTTTCTGCTATCGGCAATATGAGCTTTTATGCGACCTGTAATTAACCAAGCTGAT 1439
QY 381 AlaLysAenAlaIleThrValGlyAlaThrGlnAenLeuArgProSerPheGlySerTyr 400
DB 1440 CAGATGATGTGATGTGAGTGGGTGCGCATTTGAC-----TTT 1475

QY 401 AlaAapAenIleAenHisValAlaGlnPheSerSerArgGly----- 414
DB 1476 GAAGATTAACATC-----GCCCGCTTTCTTCCAGGGGAATGACTTACCTGGAACTA 1526
QY 415 ProThrLysAapGlyAlaArgIleLysProApyValMetAlaProGlyThrTyrIleLeuSer 434
DB 1527 CCAGAGGCTATGTCGCGCGAAGAACTGACATTTGTC-----ACCTATGGCGCGGA 1577
QY 435 AlaArgSerSerLeuAlaProAapSerSerPheThrAlaAenHisAapSerLysTyrAla 454
DB 1578 GTCCGGGCTTCCCGTGTGAAGGGGGCTGC-----CGG 1610
QY 455 TyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAenValAlaGlnLeuArg 474
DB 1611 GCACCTCAGGAGACAGATGTCCTTCCCAAGTGTGCTGGGCTGACCTTGTATGTA 1670
QY 475 GlnHisPheValLysAenArgGlyIleThrProLysProSerLeuLeuValAlaLeu 494
DB 1671 AGCAGATGAGAAAGGGAGCTAGTG-----AATCTGCAAGTGTGAAGCAAGCCCTG 1724
QY 495 IleAlaGlyAlaAlaApyValGlyLeuGlyTyrProAenGlyAen-----GlnGly 511
DB 1725 ATTCATCAAGCCCGAGGCTT-----CTGTGTAAATGTTGAGAGAGGC 1772
QY 512 TrrGlyArgValThrLeuApyLysSerLeuAenVal-----AlaTyrValAenGlnLys 529
DB 1773 CATGGCAAGTGAATCTGCTGCGAGCTTACATGATCTCAGCAGCTACCAACCAAGCGC 1832
QY 530 SerAlaLeuSerThrSerGlnLysAlaThrTyrThrPheThrAlaThrAlaGlyLysPro 549
DB 1833 AGCTTG-----AGTCT 1844
QY 550 LeuLysIleSerLeuVal-----TrrSerAapAlaProAaSer 562
DB 1845 AGCTACATCAGCTGATGATGCTGATGCTCTCATGTCGCTTACTGTTCTCAGCCATCTAC 1904
QY 563 ThrThrAlaSerValThrLeuValAenApyLeuApyLeuValIleThrAlaProAenGly 582
DB 1905 TATGAGAAATGCCAACAATGTTAAAT-----GTCCACATCTCAATGAGC 1949
QY 583 ThrArgTyrValGlyAenApyPheSerAlaProPheAapAenThr----- 598
DB 1950 ATGCGAGTCAAGGAAATTTGATTAAGCT-----GAGTGGCGGCTTATTTA 2000
QY 599 ---AapGlyArgAenAenValGlnAenValPheIleAenSer-----ProGln 613
DB 2001 CCACAAATGAGACAAATTTGAAGTGGCTTCTCTACTCTCAGTGTATGCGCTTGG 2060
QY 614 SerGlyThrTyrThrIleGlnVal 621
DB 2061 TCAGGCTACCTGGCCATCTCCAT 2084

RESULT 10
US-11-128-049-679
Sequence 679, Application US/11128049
Publication No. US20060010513A1
GENERAL INFORMATION:
APPLICANT: Melville, Mark W.
APPLICANT: Charlebois, Timothy S.
APPLICANT: Mounts, William M.
APPLICANT: Hann, Louane E.
APPLICANT: Sinacore, Martin S.
APPLICANT: Leonard, Mark W.
APPLICANT: Brown, Eugene L.
APPLICANT: Miller, Christopher P.
TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
TITLE OF INVENTION: MAKING AND USING SAME
FILE REFERENCE: 01997.027700
CURRENT APPLICATION NUMBER: US/11/128,049
CURRENT FILING DATE: 2005-05-11
PRIOR APPLICATION NUMBER: US 60/570,425
PRIOR FILING DATE: 2004-05-11
NUMBER OF SEQ ID NOS: 7285

```
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 679
/ LENGTH: 4199
/ TYPE: DNA
/ ORGANISM: Cricetus griseus
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (130)..(150)
/ OTHER INFORMATION: n is a, c, g, or t
/ NAME/KEY: misc feature
/ LOCATION: (4177)..(4199)
/ OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-679
```

Alignment Scores:

Pred. No.:	3,76e-14	Length:	4199
Score:	242.00	Matches:	138
Percent Similarity:	38.1%	Conservative:	86
Best Local Similarity:	23.5%	Mismatches:	234
Query Match:	7.4%	Indels:	130
DB:	14	Gaps:	27

US-10-784-870-4 (1-639) x US-11-128-049-679 (1-4199)

```
QY HsElleValGlnPheAnGlyProIleLeuGluGluThryGlnIlyLeuGluThr 106
   |||||
DB 552 TAAATGTGCTTTCACAGGATCTTTCACAGCCAAAGCTAAGAACTTATTTCAGT 611
QY 107 G1yAlaLys-----11Leu 111
   ||
DB 612 GCTCTGAAGACGAGTGAAGACATGGAATTAATCTCGAAGAACCATCTCACT 671
QY 112 AAspTyrlleProAspTyrlleValIleValIleGluIlyAspValIleSerIlyVal 131
   |||||
DB 672 GACTACCTGAGTGAATTTGAG--GTGATTCAGATTAAGAGAGAGAGAGAGCCGGGCTG 728
QY 132 ArgSerIleGluHleValIleGluSerValIleProTyrlleProIlySerIleAspPro 151
   |||||
DB 729 CTCACACTGGAAGATATCTCAACATCAAGCCGGGTACCTTCACAGCAAGCTTTTCT 788
QY 152 GlnLeuPheThryGlyValIleSerThryLeuValIlyValIleLeuAlaLeuAspThryGln 171
   |||||
DB 789 TCCCTGAAGTTGCTGATCTGACCCCATTTGGCCA-----TGTAAAGAACTCGGTG 842
QY 172 AsnAsnIlyGluValIleLeuArgIlyIleGluGluIleAlaGlnTyrlleValIleSerIly 191
   |||||
DB 843 AGCCAGAGTGCAGATCATCAAGACCCCTGAGAGAGAGCCAGTCTCCCTGGGCTCTGGA 902
QY 192 AspValIleThryIleThryAlaIlySerProGluTyrlleValIleMetAspAspValIleArgIly 211
   |||||
DB 903 TTCCTGAGATGCAAGAGAGAGATTCAGCCGGCATCTCTAGAGCCATTCCTGAGACG 962
QY 212 Ile-----ValIlyAlaAspValIleAlaGlnSerTyrlleTyrlleGlnIly 228
   |||||
DB 963 GTTGCCAGACATCTGAGGAGATGCTGTGG--CAGATGGAGATACAGAGTCTCAT 1019
QY 229 GlnIleValIleAlaIleAspThryIlyLeuAspThryIlyArgAsnAspSerMetIly 248
   |||||
DB 1020 GTCAGAGTGTGCTTTTGTATCTGAGGCTC-----AGTGAGAGCAT 1061
QY 249 GluAlaPheArgGlyIlySerIleThryAlaIleuTyrlleValIleArgThryAsnAsnAlaAsn 268
   |||||
DB 1062 CCACACTTCAGAGAT-----GTGAAGAGAGAGAACCACTGGAGCAAT 1103
QY 269 -----AspThryAsnGlyIlySerIlyThryIleValIleAlaGlnTyrlleValIleuGly 283
   |||||
DB 1104 GAGCGAGACCTGATGATGAGCTGGGCTGAGCATTTGTGCGAGGT--GTGATTC 1160
QY 284 AsnGlyAlaThryAsnIlyGlyMetIleProGlnAlaAsnLeu--ValPheGlnSerIle 302
   |||||
DB 1161 AGCATGAGAGAGTGCAGAGATTTTGCCTCAGATGAGAGTGCAGATCTTCGGGTCTTT 1220
```

```
QY 303 MetAspSerSerGlyIlyLeuGlyIlyLeuProSerAsnLeuGlnThryLeuPheSerGln 322
   |||||
DB 1221 ACCAACAATCAG-----GTGTCTTCACATCTGTGTTTGGAC 1259
QY 322 AlaPheSerAlaGlyAlaArgIleHleThryAsnSerTyrlleValIleAlaValAsnGlyAla 342
   |||||
DB 1260 GCTTTCACATGATCCATCTTAAGAGATGATGATTTCTTAACCTTAAGCATCGCGAGCT 1319
QY 343 TyrlleThryAspSerArgAsnValAspAspTyrlle-----ValArgIlyAsnAspMetThry 360
   |||||
DB 1320 GACTTCATGATCATCTCTTGTGTTGACAGAGGTGGAGATTAACGCTTAACATGATATC 1379
QY 361 IleLeuPheAlaIleGlyAsnGluArgProAsnGlyIlyThryIleSerAlaProGlyThry 380
   |||||
DB 1380 ATGTGTTTGTCTGAGATGATGATGATGATCTTTATGCTGACCTGAAATTAACCGAGCTGAT 1439
QY 381 AlaIlyAsnAlaIleThryValIleValIleThryGlnLeuArgProSerPheGlySerTyrlle 400
   |||||
DB 1440 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
QY 401 AlaAspAsnIleAsnHleValIleAlaGlnPheSerSerArgGly-----414
   |||||
DB 1476 GAAAGTAACTC-----GCCGCTTTCTTCAGGGGAGAGATGATGATGATGATGATGAT 1526
QY 415 ProThryIlyAspGlyIlyArgIleIlyProAspValMetAlaProGlyThryIleLeuSer 434
   |||||
DB 1527 CCAGAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1577
QY 435 AlaArgSerSerLeuAlaProAspSerPheThryAlaAsnHleAspSerTyrlleVal 454
   |||||
DB 1578 GTGCGGGTCTCGGTGTAAGGGGGCTGC-----CGG 1610
QY 455 TyrlleGlyIlyThrySerMetAlaThryProIleValAlaGlnAsnValAlaGlnLeuArg 474
   |||||
DB 1611 GCACCTTCAGAGACAGATGCTGCTCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1670
QY 475 GluHlePheValIlyAsnArgGlyIleThryProIlyProSerIleLeuValAlaIleu 494
   |||||
DB 1671 AGCACAGTGCAGAGAGGAGAGCTAGT-----AATCTGCTGAGTGAAGAGAGCCCTG 1724
QY 495 IleAlaGlyAlaAlaAspValIlyLeuGlyTyrlleProAsnIlyAsn-----GlnGly 511
   |||||
DB 1725 ATTCATGATGAGCCGAGGCTT-----CTGTGTATTAATGATTCAGAGCAAGCC 1772
QY 512 TyrlleArgValIlyThryLeuAspIlySerLeuAsnVal-----AlaTyrlleAsnGlnSer 529
   |||||
DB 1773 CATGGCAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1832
QY 530 SerAlaLeuSerThrySerGlnIlyAlaThryThryThryThryAlaGlyIlyPro 549
   |||||
DB 1833 AGCTTG-----AGTCT 1844
QY 550 LeuIlyIleSerLeuVal-----TrpSerAlaProAlaSer 562
   |||||
DB 1845 AGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
QY 563 ThryThryIleSerValIlyThryLeuAsnAspLeuValIleThryAlaProAsnGly 582
   |||||
DB 1905 TATGAGAGATGCTCAACAAATTTGTTAAT-----GTCAACATCTCAATATGGC 1949
QY 583 ThryArgTyrlleGlyAsnAspPheSerAlaProPheAspAsnAsnTrp-----598
   |||||
DB 1950 ATGGAGATCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2000
QY 599 ---AspIlyArgAsnAsnValIlyAsnValIleAsnSer-----ProGln 613
   |||||
DB 2001 CCACAGAGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2060
QY 614 SerGlyThryTyrlleGluVal 621
   |||||
DB 2061 TCAAGGTAACCTGCTCATCTTCAT 2084
```

RESULT 11

US-10-510-386-217
; Sequence 217, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groch
; APPLICANT: Clausen, Steen Troels
; APPLICANT: Oleen, Peter Blarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510.386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 217
; LENGTH: 5296
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(4796)
US-10-510-386-217
Alignment Scores:
Pred. No.: 1,14e-13 Length: 5296
Score: 238.50 Matches: 163
Percent Similarity: 34.7% Conservative: 95
Best Local Similarity: 21.9% Mismatches: 248
Query Match: 7.3% Indels: 237
DB: 8 Gaps: 36
US-10-784-870-4 (1-639) x US-10-510-386-217 (1-5296)
QY 3 LysLysLysValPheLeuSerValLeuSerAlaAlaIle-----LeuSerThrVal 20
DB 516 AGAAAAAGCGTTTCCACCATTTAAGCGTTTGAATCGCGCTCGCTTCATGCC 575
QY 21 ALLeuLysAsnProSerAlaGlyAspAlaArgThrPheAspLeuAsp----- 36
DB 576 GCGGTTTCAGATGCCGCGGGAAGCCCGCACTTACCTTCAGAAAGAAAGCAGCCGCA 635
QY 37 PheLysGlyIleGlnThrThrThrAspValSerGlyPheSerLysGlnArgGlnThrGly 56
DB 636 GGGAAAGGGAATTTCAAAAACCTCTTGTCAACAGTTCAAAAAGAAAGATCA----- 689
QY 57 AlaAlaAlaPheLeuValGluSerGluAsnValLysLeuLeuLysGlyLeuLysLys 76
DB 690 ---GTCACTTTTGAATTAAGTGAAGATCAAGTCAATACCAAAAGTGGCCAAACAG 746
QY 77 LeuGluThrValProAlaAsnAsnLysLeuHisIleVal-----GlnPheAsn 92
DB 747 GCGCAAAAGAACCGAAGAAAGTCCGTAGCGCTCTAAACAGAAATCAAAAAGCG 806
QY 93 GlyProIleLeuGluGluThrLysGlnLysLeuGluThrThrGlyAlaLysIleLeuAsp 112
DB 807 TCGGCTGTGTTCCTTAAAGATCAAGCGATGATCGCAAAAGCATCTGAAACGT 866
QY 113 TyrIleProAspTyrAlaTyrIleValGluTyrGluGlyAspValGlnSerLysValArg 132
DB 867 TATTTAAAAAAGCAG-----GAAAAACAGGGGACGTAAAA---AAATCAGA 911
QY 133 Ser-----IleGluHisVal 137
DB 912 TCTTATTAATATGTCAACGCGCATGCGCTCATGCGAAGAGATCATGAAAGTA 971
QY 138 GluSer-----ValGluProTyrLysProLysTyrLysIleAspProGlnLeuPhe 154
DB 972 GCGGCTTTTCCGAAGTGAAGAAAGTCTTCCAAACGAAAA----- 1013
QY 155 ThrLysGlyAlaSerThrLeuValLysAlaLeuAlaLeuAspThrLysGlnAsnAsnLys 174
DB 1014 ---AGGCACGTGATCAACCGACTAAACGCTGCTGTGAAA 1049

QY 175 GluValGlnLeuArgGlyIleGluGlnIleAlaGlnTyrValAlaSerAsnAspValHis 194
DB 1050 AAATCAGACGACCGCAAGTGAAGAAATCGAATG-----AAATCAAT 1094
QY 195 TyrIleThrAlaLysProGluTyrLysValMetAsnAspValAlaArgGlyIleValLys 214
DB 1095 CGAGTCACGCAACCAAAAGCTGGAATTC----- 1124
QY 215 AlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnGlyGlnIleValAlaValAla 234
DB 1125 -----GATATGACGATCGGGCAGCGTTGTGCGTCGAT 1160
QY 235 AspThrGlyLysLeuAspThrGlyLysAsnAspSerSerMetHisGluAlaPheArgGly--- 253
DB 1161 GATACGGCGGTGACGTGG-----GATCATCCGGCTTGAAGAAATACCGCGGATAT 1214
QY 254 -----LysIleThrAlaLeuTyrAlaLeuGlyArg 263
DB 1215 GACCCGACGATCCCATTCAGCTTACCGAATTCAGTTGTATGATGCGGTTTCAGCC 1274
QY 264 ThrAsnAsnAlaAsnAspThrAsnGlyHisGlyThrHisValAlaGlySerValLeuGly 283
DB 1275 GCTTGTGAACCATTTGACATCTGAGACGACGACGATGTAACCGGACGATGTCGCGC 1334
QY 284 -----AsnGlyAlaThrAsnLysGlyMetAlaProGlnAlaAsnLeuValPheGln 300
DB 1335 TCAGAGGCTCAGCGCCCAAAACCAATTCGGCGTGGCGCGGATGATGTCGTG 1394
QY 301 SerIleMetAspSerSerGlyLysLeuGlyGlyLeuProSerAsnLeuGlnThrLeuPhe 320
DB 1395 AAAGCTTTTTCAGATGACGGCGGTACCGAC----- 1424
QY 321 SerGlnAlaPheSerAlaGlyAlaArgIle----- 330
DB 1425 GAAGACCTTATTCAGACGATGATGATTTTGGCGCAAAAGTAAAGACGCAACCG 1484
QY 331 His-----ThrAsnSerTrp-----GlyAlaAlaValAsn 340
DB 1485 CACCTGAAATGCGCGCGATGTCGTAATTAATCATAGTGCAGAGGAGACCGGATGAT 1544
QY 341 GlyAlaIleThrThrAspSerArgAsnValAspAspTyrValArgLysAsnAspMetThr 360
DB 1545 GAGCTTTTAC-----AGGATATCGTAAAGCTGGCGAGCTGCGGCATCTTT 1592
QY 361 IleLeuPheAlaIleGlyAsn-----GluArgProAsnGly---GlyThrIleSer 376
DB 1593 CTGAGATTCCTCGGGGAACGTGATTCGCCAATCCGGAGAGACCTGCTTCGATTCGC 1652
QY 377 AlaProGlyThrAlaLysAsnAlaIleThrValGlyAlaThrGluAsnLeuArgProSer 396
DB 1653 AATCCCGCAAACTAACCTTGAAGCGTTTCGCAACCGGACGCA----- 1694
QY 397 PheGlySerTyrAlaAspAsnIleAsnHisValAlaGlnPheSerSerArgGlyProThr 416
DB 1695 -----ACATCATTAATCGCTTACCGGATTTTCACTTCAAGGCTCTTC 1739
QY 417 LysAspGlyArgIleLysProAspValMetAlaProGlyThrTyrIleLeuSerAlaArg 436
DB 1740 CCGTATGATGACGATTAAGCTGAATATCGGCGCTGCTGTCAACATT-----CGC 1790
QY 437 SerSerLeuAlaProAspSerSerPheTyrAlaAsnHisAspSerLysTyrAlaTyrMet 456
DB 1791 TCATCTGTG---CTGGAAGCGGCTATCAAGACGCTGGAC----- 1829
QY 457 GlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGlnLeuArgGluHis 476
DB 1830 ---GGCACTTCATATGCTGCGCGCATGTGCGAGCGCGCGCTCTTCAAGCAAGCC 1886
QY 477 Phe-----Val 478
DB 1887 GATTCGTGATCATCTTGTATGATGACGAAAGATTTTGTATGAAACTGCAACACCGCTC 1946

QY 479 LysAsnArgGlyLeuThrProLysProSer-----LeuLeuLysAla 492
Db 1947 ACGAGACGCAATTATTCGATCGCGAACACGATACGCCACGACCTCGTAACGTG 2006
QY 493 AlaLeuLeuAlaGlyAlaAspValGlyLeuGlyTyrProAsnGlyAsnGlyTyr 512
Db 2007 TTGACGACGATGTCGCTGACAGACGGGTTGGCAGACGGAGACGAAAGTC----- 2060
QY 513 GlyArgValThrLeuAspLysSer-----LeuAsnValAlaTyrValAsnGlyLys 529
Db 2061 GGAAGAACGACGACGAAAGATCGCGCGCTTAAGCATCAAGAGATCACCGAA--- 2117
QY 530 SerAlaLeuSerThrSerGlyLysAlaThrTyrThrPheThrAlaThrAlaGlyLysPro 549
Db 2118 -----ATTACTCAGGGCTCGAAGC-----CCT 2141
QY 550 LeuLys-----LleSerLeuVal----- 555
Db 2142 CTTAAAGCTGATGTAGAACAGATGTCAGCATCATTCGGTCAAAATTGACCTATAAAC 2201
QY 556 -----TrpSerAspAlaProAlaSerThrThrAlaSerValThrLeuValAsn 571
Db 2202 GATGACGCGCATGTGGAGACGATCGCTGCCAAACAACTTCGCGCATATAAAAG 2261
QY 572 AspLeuAspLeuValIleThrAlaProAsnGlyThr-----ArgTyrValGly 587
Db 2262 ACCTACGAGCGCGCTGACTGTTCTGGAAGGAAACCTTTCCTACAGATGATGTC 2321
QY 588 AsnAspPheSerAlaProPheAspAsnAsnTyrAspGlyArgAsnAsnValGluAsnVal 607
Db 2322 ACCGATCTC-----GAGGCAACAAACGAAATCC--- 2351
QY 608 PheIleAsnSerProGlnSerGlyThrTyrThrIleGluValGlnAlaTyrAsnValPro 627
Db 2352 -----AAAGTGTACGAAGTCCCG 2369
QY 628 ValGlyPro 630
Db 2370 ATTTCCCG 2378
RESULT 12
US-11-156-062-13
; Sequence 13, Application US/1156062
; Publication No. US20050281773A1
; GENERAL INFORMATION:
; APPLICANT: Wieland, Susanne
; APPLICANT: Polanyi-Bald, Laura
; APPLICANT: Prueser, Inken
; APPLICANT: Stehr, Regina
; APPLICANT: Maurer, Karl-Heinz
; TITLE OF INVENTION: SUPPLISIN VARIANTS, WITH IMPROVED PEPTIDOLASE ACTIVITY
; FILE REFERENCE: HENK-0134 / H698
; CURRENT APPLICATION NUMBER: US/11/156, 062
; PRIOR FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/014127
; PRIOR FILING DATE: 2003-12-20
; PRIOR APPLICATION NUMBER: DE 102 60 903.9
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 824
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Variant T58A/L216W of SuplilisIn Carlsberg
; NAME/KEY: CDS
; LOCATION: (1)..(822)
US-11-156-062-13
Alignment Scores: 2.09e-13 Length: 824
Pred. No.:

Score: 226.50
Percent Similarity: 43.1%
Best Local Similarity: 30.6%
Query Match: 6.9%
DB: 14
Gaps: 11
US-10-784-870-4 (1-639) x US-11-156-062-13 (1-824)
QY 212 IleValIleAlaAspValAlaGlnSerSerTyrGlyLeuTyrGlyGlnIleVal 231
Db 28 CTCATTAAGCGGACCAAGTGCAGGCTCAA---GGCTTAAGGACGCAATGTAAAGTA 84
QY 232 AlaValAlaAspThrGlyLeuAspThrGlyArgAsnAspSerSerMetIleGluAlaPhe 251
Db 85 GCGGTCTCGAATACAGAAATCCAGCTTCATCCGACTTGAACCTATCCGCGGA--- 141
QY 252 ArgGlyLysIleThrAlaLeuTyrAlaLeuGlyArgThrAsnAsnAlaAsnAspThrAsn 271
Db 142 -----GCAAGCTTGTGGCTGGCGAAAGCTTATTAAGCC---GACGCAAC 183
QY 272 GlyAsnGlyThrIleValAlaGlySerVal-----LeuGlyAsnGlyAlaThrAsnLys 289
Db 184 GGAACGCGACACATGTTCCGGTACAGTACGTGCGCTTGACAAATCAACGGGTGTTTA 243
QY 290 GlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyLeu 309
Db 244 GCGGTTCGCGCAAGCGTATCTGTACCGGTTAAAGTACATAATTCAGCGGA----- 297
QY 310 GlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArg 329
Db 298 ACGGATCATACAGCGGCACTTTGAACGGAATCAAGTGGCGCAACAAACGCAATGAT 357
QY 330 IleIleThrAsnSerTrpGlyAlaAlaValAsnGlyAlaTyrThrAspSerArgAsn 349
Db 358 GTTATCAATATGAGCTTGGCGGACCA-----TCAGGCTCAACAGCGATGAACAG 408
QY 350 ValAspAspTyrValArgLysAsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArg 369
Db 409 GCAATGCAATGCAATGCAATGCAAGAGGCTGCTGTGAGCTGCAGCAGGAAACGCGGA 468
QY 370 ProAsnGly-----GlyThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrVal 387
Db 469 TCTTCAGAAACACGAATACATTTGGCTATCTCGGAATACGATTCCTCATTCCTGTT 528
QY 388 GlyAlaThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsnIleVal 407
Db 529 GGTGCG-----GTAGACTTAAACAGCAAGA 555
QY 408 AlaGlnPheSerSerArgGlyProThrLysAspGlyArgIleLysProAspValMetAla 427
Db 556 GCTTCATTTCCAGGCTCGA-----GCAGACTTGAAGATCATGCT 597
QY 428 ProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheTrpAla 447
Db 598 CTTGCGCAGCGGCTATACAGCACTTAACCAACG----- 630
QY 448 AsnIleAspSerLysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAla 467
Db 631 -----AACCTTAATGCAACATGGAACGGAAGTCAATGGCTTCCTCATGTAGG 681
QY 468 GlyAsnValAla-----GlnLeuArgGluIlePheValLysAsn 480
Db 682 GAGGACGACGCTTTGATCTTGCAAAACATCGAACTTTACGCTTCAACAAATCGGCAAC 741
QY 481 Arg 481
Db 742 CGT 744
RESULT 13
US-11-065-943-59
; Sequence 59, Application US/11065943
; Publication No. US20050250131A1
; GENERAL INFORMATION:

APPLICANT: JESTIN, JEAN-LUC
APPLICANT: VICHIER-GUERRE, SOPHIE
APPLICANT: PERRIS, STEPHANE
TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES,
TITLE OF INVENTION: METHODS FOR OBTAINING THE SAME, AND APPLICATIONS OF THE SAME
FILE REFERENCE: 266426USOXCIP
CURRENT APPLICATION NUMBER: US/11/065,943
PRIOR FILING DATE: 2005-02-25
PRIOR APPLICATION NUMBER: US 10/787,219
NUMBER OF SEQ ID NOS: 106
SOFTWARE: Patent in version 3.3
SEQ ID NO 59
LENGTH: 825
TYPE: DNA
ORGANISM: Bacillus amyloliquefaciens
US-11-065-943-59

Alignment Scores:

Pred. No.:	2,09e-13	Length:	825
Score:	226.50	Matches:	82
Percent Similarity:	39.4%	Conservative:	29
Best Local Similarity:	29.1%	Mismatches:	88
Query Match:	6.9%	Indels:	83
DB:	14	Gaps:	10

US-10-784-870-4 (1-639) x US-11-065-943-59 (1-825)

213 VallyValAAspValAaInserSerTyrglyLeuTyrglyGlnGlyGlnIleValAla 232
:::|||||
31 ATTAAAGCCCT---GCTTGCACTCTCAAGGCTACACTGATCAATGTAAAGTAGCC 87
:::|||||
233 ValAlaAspThrglyLeuAspThrglyArGAsnAspSerSerMetHicglAlaIleApeArg 252
|||:::|||||
88 GTTATCGACAGCGGATTCGAT-----TCTTCATCTCGATTAAAG 129
:::|||||
253 GlyIleIleThraIleuTyrglyAlaLeuGlyArGThraAsnAlaAsnAspThraGly 272
|||:::|||||
130 GTAGCAGCGGAGCCGACATGCTGTTCTTGAAACAAATCTTCCAGACACAACTCT 189
|||:::|||||
273 HicglYThrHicglValAlaGlySerVal-----LeuGlyAsnGlyAlaThraAsnGly 290
|||:::|||||
190 CACGGAACCTCACGCTTCCGACACAGTTGCGGCTTTAAATCAATCGGTGATTAGGC 249
|||:::|||||
291 MetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerMetHicglAlaIleApeArg 310
|||:::|||||
250 GTTGCGCCAGCGCATCTTACGCTGTAAAGTTCTCGGTCGACGCT----- 300
|||:::|||||
311 GlyLeuProSerAsnLeuGlnThraLeuPheSerGlnAlaPheSerAlaGlyAlaArgIle 330
|||:::|||||
300 ----- 300
331 HicThraAsnSerTrpGlyAlaAlaValaAsnGlyAlaTyTrThraAspSerArgAsnVal 350
|||:::|||||
301 ---TCCGGCCATACAGCTGATCATTTAAAGGATGAGTGGCGCATCGAAACAAATATG 357
|||:::|||||
351 AAsp----- 351
358 GACGTTATTAAACATGAGCCTCGCGGACCTTCGTTCTGCTGCTTTAAAGCGGACATT 417
|||:::|||||
352 AspTyTrValAlaArgLeuAsnAspMetThrIleLeuPheAlaAlaGlyAsnGlyArgProAsn 371
|||:::|||||
418 GATTAAGCCGTTGACATCCGCGCTGAGTCGTTGCGCGACGCCGTTAAACGAGCACTTCC 477
|||:::|||||
372 Gly-----GlyThrIleSerAlaProGlyThraIleAsnAlaIleThraValGlyAla 389
|||:::|||||
478 GGCAGCTCAAGCAGCAGTGGGCTACCTGGAATACCTTCCTGCAATGCAAGTAGCGCT 537
|||:::|||||
390 ThrgIAsnLeuArgProSerPheGlySerTyrglyAlaAspAsnIleAsnHicglAlaGln 409
|||:::|||||
538 -----GTTGACAGCAGCAACCAAAAGACATCT 564
|||:::|||||

410 PheSerSerArgIleProThraIleAspGlyArgIleTyProAspValMetAlaProGly 429
|||||
565 TTCGAGCGGTAGGACT-----GAGCTTATGTCATGCGACCTGCG 606
|||||
430 ThrTyrlleuSerAlaArgSerSerLeuAlaProAspSerPheTrpAlaAsnHic 449
:::|||||
607 -----GTACTATCCAAAGCAGCTTCC----- 630
|||||
450 AspSerTyTrAlaTyTrMetGlyTyTrSerMetAlaThraProIleValAlaGlyAsn 469
:::|||||
631 GGAACAAATATACGGCGCGTACCAAGGTACATGCAATGCAATCTCCGACGTTGCCGAGCG 650
|||||
470 ValAla 471
|||
691 GCTGCT 696
|||
RESULT 14
US-11-156-062-11
Sequence 11, Application US/11156062
PUBLICATION NO. US20050281773A1
GENERAL INFORMATION:
APPLICANT: Wieland, Susanne
APPLICANT: Polanyi-Bald, Laura
APPLICANT: Prueser, Inken
APPLICANT: Stehr, Regina
APPLICANT: Maier, Karl-Heinz
TITLE OF INVENTION: SUBSTITUTED VARIANTS, WITH IMPROVED PERHYDROLASE ACTIVITY
FILE REFERENCE: HENK-0134 / H5698
CURRENT APPLICATION NUMBER: US/11/156,062
PRIOR FILING DATE: 2005-06-17
PRIOR APPLICATION NUMBER: PCT/EP2003/014127
PRIOR FILING DATE: 2003-12-20
PRIOR APPLICATION NUMBER: DE 102 60 903.9
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent in version 3.3
SEQ ID NO 11
LENGTH: 824
TYPE: DNA
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Variant T58A/L89S/N96D/L216W/N217D of Subtilisin Carlsberg
NAME/KEY: CDS
LOCATION: (1)..(822)
US-11-156-062-11

Alignment Scores:

Pred. No.:	2.65e-13	Length:	824
Score:	225.50	Matches:	86
Percent Similarity:	42.7%	Conservative:	34
Best Local Similarity:	30.6%	Mismatches:	108
Query Match:	6.9%	Indels:	53
DB:	14	Gaps:	11

US-10-784-870-4 (1-639) x US-11-156-062-11 (1-824)

212 IleVallyValAAspValAaInserSerTyrglyLeuTyrglyGlnGlyGlnIleVal 231
:::|||||
28 CTCATTAAAGCGGACAAATGCAAGCTCA---GCTTTAAAGGACGGAATGTAAAGTA 84
|||:::|||||
232 AlaValAlaAspThrglyLeuAspThrglyArGAsnAspSerSerMetHicglAlaIleApe 251
|||:::|||||
85 GCCGCTCTGATTAACAGATCCAGGCTTCTCATCCGACCTTGAACGTAGTCGCGGCA--- 141
|||:::|||||
252 ArgGlyIleIleThraIleuTyrglyAlaLeuGlyArGThraAsnAlaAsnAspThraAsn 271
|||:::|||||
142 -----GCAAGCTTGTGCTGCGGAGACTTATTAAGCC---GACGCGCAAC 183
|||:::|||||
272 GlyHicglYThrHicglValAlaGlySerVal-----LeuGlyAsnGlyAlaThraAsnGly 289
|||:::|||||
184 GGAACGCGCACATGTTGCCGTTACAGTAGCTGCGCTGACACATTAACAGGCTGTTTAA 243
|||:::|||||

```

Oy 290 GlyMetAlaProGlnAlaAsnLeuValPheGlnSerIleMetAspSerSerGlyGlyLeu 309
Db 244 GCGCTGGCGCCAGAGCGTATCTCTGTAACGGGTTAAAGTATGATTCAGCCGA----- 297
Oy 310 GlyGlyLeuProSerAsnLeuGlnThrLeuPheSerGlnAlaPheSerAlaGlyAlaArg 329
Db 298 AGCGATCATACAGCGGCACTTGTAAACGGAAATCGATGGCGCAACAAACGGCATGAT 357
Oy 330 IleIsthrAsnSerTrpGlyAlaAlaValAsnGlyAlaTyThrThrAspSerArgAsn 349
Db 358 GTTATCATATATAGACCTTGCGGGAGCA-----TCAGGCTCGCAGACGATGAAACAG 408
Oy 350 ValAspAspTyrValArgIysAsnAspMetThrIleLeuPheIleAlaGlyAsnGluArg 359
Db 409 GCAAGTCGACMATGCAATATGCGAGAGGGGCTGCTGTGTAGCTGCACAGGAAACAGCGGA 468
Oy 370 ProAsnGly-----GlyThrIleSerAlaProGlyThrAlaIysAsnAlaIleThrVal 387
Db 469 TCTTCAGGAAACACGAATACATTTGGCTATCTTCGCGAAATACGATTTCTGCATCGCTGT 528
Oy 388 GlyAlaPThrGluAsnLeuArgProSerPheGlySerTyrAlaAspAsnIleAsnIleVal 407
Db 529 GGTGGG-----GTAGACTCTMAACAGCAACAGA 555
Oy 408 AlaGlnPheSerSerArgGlyProThrIysAspGlyArgIleIysProAspValMetAla 427
Db 556 GCTTCATTTTCCAGCGCTCGGA-----GCAAGAGCTTGAAAGTCATGGCT 597
Oy 428 ProGlyThrTyrIleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheTrpAla 447
Db 598 CTTGGCGCAGGCGGTATACAGCACTTACCCACAG----- 630
Oy 448 AsnHisAspSerIysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAla 467
Db 631 -----AACACTATACACATGAAACGGAACGTCATGGCTTCTCTCATGTAGCG 681
Oy 468 GlyAsnValAla-----GlnLeuArgGluHisIspValIysAsn 480
Db 682 GGAGCAGCAGCTTGTATCTTGTCAAACATCCGAACCTTTCACCTTCAACAAGTCGCGAAC 741
Oy 481 Arg 481
Db 742 CGT 744

RESULT 15
US-11-008-331-1/c
; Sequence 1, Application US/11008331
; Publication No. US20050244925A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC386-PCT
; CURRENT APPLICATION NUMBER: US/11/008,331
; PRIORITY FILING DATE: 2004-12-08
; PRIOR APPLICATION NUMBER: EP9719637.2
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2760 bp
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-11-008-331-1

```

Qy 503 leuGIyTYrProAengIyAengIingIyTrpGIyArGValFhrLeuAapLySerLeuAsn 522
 Db 1684 -----TGGAAAAGGTTAATCAACGTACAGCAGCTGC 1652
 Qy 523 ValAlaTYrValAengIuSerSerAlaLeuSerThr 534
 Db 1651 ACAATATAGTAAAAAGAGCAGCGTCTCCATACC 1616

Search completed: April 8, 2006, 08:55:20
 Job time : 2743 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:56:02 ; Search time 9420.85 Seconds

(without alignments)
11602.984 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923
Sequence: 1 atcgagaagaagaagaagc.....cgttgcaatgtgtaataa 1923

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pr:.*
9: gb_pro:.*
10: gb_str:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_hlg:.*
15: gb_pl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1923	100.0	1923	1	AB051423	AB051423 Bacillus
2	1923	100.0	1923	6	CQ888285	CQ888285 Sequence
3	1923	100.0	1923	6	CQ889309	CQ889309 Sequence
4	1923	100.0	1923	6	AR368117	AR368117 Sequence
5	1923	100.0	1923	6	AR562465	AR562465 Sequence
6	1919.8	99.8	1923	6	AR368118	AR368118 Sequence
7	1919.8	99.8	1923	6	AR562466	AR562466 Sequence
8	1887.8	98.2	1923	1	AB084155	AB084155 Bacillus
9	1477.2	76.8	1920	1	AB046403	AB046403 Bacillus
10	1477.2	76.8	1920	6	AR562464	AR562464 Sequence
11	1477.2	76.8	1920	6	AR562464	AR562464 Sequence
12	1305	67.9	1305	6	AX839476	AX839476 Sequence
13	1166.4	60.7	3003	6	AR069954	AR069954 Sequence
14	1166.4	60.7	3003	6	BD062155	BD062155 Nucleic a
15	936.2	48.7	2218	6	E03808	E03808 DNA encodin
16	854.6	44.4	1302	1	AB046406	AB046406 Bacillus
17	740.4	38.5	1299	1	AB046405	AB046405 Bacillus
18	738.8	38.4	1299	1	AB046402	AB046402 Bacillus

19	737.2	38.3	1299	1	AB046404	AB046404 Bacillus
20	81	4.2	346148	1	BX842652	BX842652 Bdeliowib
21	54	2.8	7218	6	I66494	I66494 Sequence 14
22	53.6	2.8	2000	6	AX655393	AX655393 Sequence
23	50	2.6	2000	6	AX655393	AX655393 Sequence
24	48	2.5	54	6	AX839494	AX839494 Sequence
25	47.2	2.5	5890	2	DDU60086	DDU60086 Dictyosteli
26	46.6	2.4	1239	1	AF305633	AF305633 Thermoana
27	46.6	2.4	1239	1	AY028704	AY028704 Thermoana
28	45.2	2.4	40499	8	AC134821	AC134821 Homo sapi
29	45.2	2.4	168242	8	AC140830	AC140830 Homo sapi
30	45.2	2.4	173275	14	AC145031	AC145031 Homo sapi
31	45.2	2.3	1415	6	AX566129	AX566129 Sequence
32	45	2.3	12295	6	AE013026	AE013026 Thermoana
33	45	2.3	16831	8	HS523C21	HS523C21 Human DNA
34	45	2.3	125020	8	AF429315	AF429315 Homo sapi
35	44.8	2.3	154712	14	AC139734	AC139734 Sus acrofe
36	44.6	2.3	176462	8	AC159827	AC159827 Pan trogl
37	44	2.3	12384	8	AC116347	AC116347 Homo sapi
38	43.6	2.3	168430	14	AC145037	AC145037 Homo sapi
39	43.6	2.3	175559	14	AC145101	AC145101 Homo sapi
40	43.4	2.3	110000	15	CR382122_08	CR382122_08 Continuation (9 of
41	43.4	2.2	21647	14	AC114167	AC114167 Rattus no
42	43.2	2.2	250541	1	AE017004	AE017004 Bacillus
43	43	2.2	136613	14	AC164927	AC164927 Ginglymos
44	42.8	2.2	834	6	BD110490	BD110490 EST and e
45	42.4	2.2	834	6	BD110490	BD110490 EST and e

ALIGNMENTS

RESULT 1
AB051423
LOCUS
DEFINITION
AB051423
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLES
JOURNAL

AB051423 1923 bp DNA linear BCT 10-MAY-2002
Bacillus sp. KP43 PRO gene for protease, complete cds.
AB051423
AB051423.2 GI:20521154
Bacillus sp. KSM-KP43
Bacillus sp. KSM-KP43
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Itoh, S. and Saeki, K.
new protease
TITLES
JOURNAL
2. (bases 1 to 1923)
REFERENCES
AUTHORS
TITLES
JOURNAL
Direct Submission
Submitted (21-NOV-2000) Katsuhisa Saeki, KAO CORPORATION;
2606, AKABANE, ICHIRAIMACHI, HAGA, TOCHIGI 321-3486, Japan
(E-mail: 387185@kcaetnet.kao.co.jp, Tel: 81285687471 (ex. 7471),
Fax: 81285687403)

COMMENT
On May 9, 2002 this sequence version replaced gi:14164344.

FEATURES

source
1..1923
/organism="Bacillus sp. KSM-KP43"
/mol_type="genomic DNA"
/strain="KP43"
/db_xref="taxon:109322"
1..1923
/gene="PROF"
1..1923
/gene="PROF"
/codon_start=1
/catalytic_site=1
/product="protease"
/protein_id="BAB55674.2"
/db_xref="GI:20521155"
/translation="MRKKKVLSTLSAAALSTVALSNPSAGANPDLDEKGIQTT
TDKGSKGGTGAALFLVESNVGLPKGLKKLETPANNLGLHIFGPKLSTGK
OLEKGAIKLIDYIPYAYIVVEYGVKSTSTIRHVESVPLPIYRIDPOLPTGAS
ELVKAVADLTQKKNKEVQLRGIEQLAQPAISNDVLITPAKPKYKMDVAKRGIVADV

AOSSYGLYGGGQIVAVADTGLDTRNDSSWHEAFRGKITLALVLRGTNNANDTNGHT
HYAGSYLGNSTKMGKAAPOANLYRQSLIMSGLGLBSLQTLFQOASARAHITN
SWKAVNGATTTDSRVNDVYRKNDITLIPASNGPNGGTISABGTAKOALTVGATN
NLRPSFGSTADNINHVAQSSRGPTDGRKPDVMAKPGFTLISARSLAPDSFWMNH
DSKRAYMGTSNATPIVAGNVQALRHPVNNRGTIPKPSLAKALJAGADLGYEN
GNGWGRVTLDEKSLNVAYNVESLSLTSOKATYSFATAGKPKLKLWMSDAPASTA
SVTLVNDLIVITAPNGTOYVNDFTSPYNDNDGRNNENVINAPQSGTYTIEVQA
YVNPVGQRFPSLAIVN

ORIGIN

Query Match	100.0%; Score 1923; DB 1; Length 1923;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGAGAAAGAAAGAAAGGTTTTTATCTGTTTTATCAGCTGCAGCATTTTGTGACT 60
DB	1	ATGAGAAAGAAAGAAAGGTTTTTATCTGTTTTATCAGCTGCAGCATTTTGTGACT 60
QY	61	GTTGCTTAAATATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120
DB	61	GTTGCTTAAATATCATCTGCAAGTGTGCAAGAAATTTTATCTGATTTCAAGGA 120
QY	121	ATTGACACACAACTGATGCTAAAGGTTCTCCAGCAGGGGAGACTGCTGCTGCT 180
DB	121	ATTGACACACAACTGATGCTAAAGGTTCTCCAGCAGGGGAGACTGCTGCTGCT 180
QY	181	TTTCTGCTGAATCTGAAAATGGAACCTCCAAAAGGTTTGAGAGAAAGCTTGAACA 240
DB	181	TTTCTGCTGAATCTGAAAATGGAACCTCCAAAAGGTTTGAGAGAAAGCTTGAACA 240
QY	241	GTCCCGGCAATATTAACCTCATTTATTCATTTACATGAGCAATTTTGAAGAAACA 300
DB	241	GTCCCGGCAATATTAACCTCATTTATTCATTTACATGAGCAATTTTGAAGAAACA 300
QY	301	AAACAGCAGCTGGAAGAAACAGGGGCAAGATTCTGACTACATACCTGATATGCTTAC 360
DB	301	AAACAGCAGCTGGAAGAAACAGGGGCAAGATTCTGACTACATACCTGATATGCTTAC 360
QY	361	ATTGTCGATATGAGGGGCGATGTTTAGTGAGCAACAGACCAATTGAGAACCTGGAATCC 420
DB	361	ATTGTCGATATGAGGGGCGATGTTTAGTGAGCAACAGACCAATTGAGAACCTGGAATCC 420
QY	421	GTGAGACCTTATTTGCGATATACAGATATAGATCCCACTTTTCAAAAAGGGGCAATCA 480
DB	421	GTGAGACCTTATTTGCGATATACAGATATAGATCCCACTTTTCAAAAAGGGGCAATCA 480
QY	481	GAGCTTGTAAAGCAGTGGCGCTTGATCAAAAGCAGAAATTAAGAGGTGCAATTAAGA 540
DB	481	GAGCTTGTAAAGCAGTGGCGCTTGATCAAAAGCAGAAATTAAGAGGTGCAATTAAGA 540
QY	541	GGCATGGAACAAATGCACAATTCGCAATTAAGCAATGATGCTATATTAACGGCAAG 600
DB	541	GGCATGGAACAAATGCACAATTCGCAATTAAGCAATGATGCTATATTAACGGCAAG 600
QY	601	CCTGAGTATAGGTGATGATGATGTTGGCGGTGGAATTTGCAAAAGCGATGTGCTCAG 660
DB	601	CCTGAGTATAGGTGATGATGATGTTGGCGGTGGAATTTGCAAAAGCGATGTGCTCAG 660
QY	661	AGCAGTACGGGTTGTATGAGCAAGGACAGATGTAAGCGGTGCGATACAGGGCTTGAT 720
DB	661	AGCAGTACGGGTTGTATGAGCAAGGACAGATGTAAGCGGTGCGATACAGGGCTTGAT 720
QY	721	ACAGGTGCAATGACAGTTGATGATGAGAGCTTCCCGGGGAAATTAACGTCATTAAT 780
DB	721	ACAGGTGCAATGACAGTTGATGATGAGAGCTTCCCGGGGAAATTAACGTCATTAAT 780
QY	781	GCATTGGAGCGGACGATATATGCAATGATGATGATGATGATGATGATGATGATGAT 840
DB	781	GCATTGGAGCGGACGATATATGCAATGATGATGATGATGATGATGATGATGATGAT 840
QY	841	TCCGTTATTAAGAAAGGCTTCCATTAATTAAGAAATGCGGCTCAGGCGCATTAAGTCTTC 900
DB	841	TCCGTTATTAAGAAAGGCTTCCATTAATTAAGAAATGCGGCTCAGGCGCATTAAGTCTTC 900

QY	901	CAATCTATATGATATGCGGTGGGGGACTTGGAGACTACCTTGCATATGCAAACTTTA 960
DB	901	CAATCTATATGATATGCGGTGGGGGACTTGGAGACTACCTTGCATATGCAAACTTTA 960
QY	961	TTACGCAAGACATACAGTGTGCTGTCAGAAATTCATACAAATCTCTGGGGACACAGAG 1020
DB	961	TTACGCAAGACATACAGTGTGCTGTCAGAAATTCATACAAATCTCTGGGGACACAGAG 1020
QY	1021	AATGGGGCTTACACAAAGATTCAGAAATGTGATGACTATGTGCGCAAAATGATATG 1080
DB	1021	AATGGGGCTTACACAAAGATTCAGAAATGTGATGACTATGTGCGCAAAATGATATG 1080
QY	1081	ACGATCTTTTGTGCTGCGGGGAATGAAGACCGAACCGGCAACATCATGTCACCAAGC 1140
DB	1081	ACGATCTTTTGTGCTGCGGGGAATGAAGACCGAACCGGCAACATCATGTCACCAAGC 1140
QY	1141	ACAGCTAAATAATGCAATTAACAGTGGAGCTACGGAACCTCCGGCCAAAGCTTTGGTCT 1200
DB	1141	ACAGCTAAATAATGCAATTAACAGTGGAGCTACGGAACCTCCGGCCAAAGCTTTGGTCT 1200
QY	1201	TATGCGGCAATATCAACATGATGACACAGTTCTTTCAAGTGAACCGCAAAAGATGGA 1260
DB	1201	TATGCGGCAATATCAACATGATGACACAGTTCTTTCAAGTGAACCGCAAAAGATGGA 1260
QY	1261	CGGATCAAAACCGGATGTGATGAGCACCGGGAAAGTTCAATCATGACGAAGATCTTCTTT 1320
DB	1261	CGGATCAAAACCGGATGTGATGAGCACCGGGAAAGTTCAATCATGACGAAGATCTTCTTT 1320
QY	1321	GCAACCGGATTCCTCTCTGCGGCGAACCATGACAGTAATATGCAATACAGGGTGAACG 1380
DB	1321	GCAACCGGATTCCTCTCTGCGGCGAACCATGACAGTAATATGCAATACAGGGTGAACG 1380
QY	1381	TCATGCTTACACCGATGCTTGTGGAAGAGTGGCACAGCTTGTGACATTTGTGAAA 1440
DB	1381	TCATGCTTACACCGATGCTTGTGGAAGAGTGGCACAGCTTGTGACATTTGTGAAA 1440
QY	1441	AAACAGGACATCACACAAAGCTTCTCTATTAAGAGGAGCAATGATGCGGCGCACT 1500
DB	1441	AAACAGGACATCACACAAAGCTTCTCTATTAAGAGGAGCAATGATGCGGCGCACT 1500
QY	1501	GACATCGGCTTGTGCTAACCGGAACGGTAAACCAAGATGGGACAGATGACATTTGGA 1560
DB	1501	GACATCGGCTTGTGCTAACCGGAACGGTAAACCAAGATGGGACAGATGACATTTGGA 1560
QY	1561	TCCTGAAAGTTCCTATGTGAACGAGTCAAGTTCTTATCCACAGCAAAAGCGACG 1620
DB	1561	TCCTGAAAGTTCCTATGTGAACGAGTCAAGTTCTTATCCACAGCAAAAGCGACG 1620
QY	1621	TACTGTTTACTGCTACCTGCGGCAAGCTTTGAAAATCTCCGTGATGATGATGATGATG 1680
DB	1621	TACTGTTTACTGCTACCTGCGGCAAGCTTTGAAAATCTCCGTGATGATGATGATGATG 1680
QY	1681	CCTGAGACCAACTGCTTCCGTAAACCTTGTCAATGATGAGACTTGTCAATTAACCGCT 1740
DB	1681	CCTGAGACCAACTGCTTCCGTAAACCTTGTCAATGATGAGACTTGTCAATTAACCGCT 1740
QY	1741	CCAAATGACACAGATATGTAAGAAATGACTTTTACTTCCGCAATCAATGATTAACG 1800
DB	1741	CCAAATGACACAGATATGTAAGAAATGACTTTTACTTCCGCAATCAATGATTAACG 1800
QY	1801	GGCGGCAATTAACGTAAGAAATGATTTATTAATGACACAAAGCGGAGACTATCAAT 1860
DB	1801	GGCGGCAATTAACGTAAGAAATGATTTATTAATGACACAAAGCGGAGACTATCAAT 1860
QY	1861	GAGGTACAGGCTTAAAGTACCGGTTGACACAGACTTCTGTTGGCAATTTGTAAT 1920
DB	1861	GAGGTACAGGCTTAAAGTACCGGTTGACACAGACTTCTGTTGGCAATTTGTAAT 1920
QY	1921	TAA 1923
DB	1921	TAA 1923

```

RESULT 2
LOCUS      CQ888285
DEFINITION Sequence 2 from Patent EP1466970.
ACCESSION  CQ888285
VERSION     CQ888285.1 GI:54304586
KEYWORDS
SOURCE      Bacillus sp. KSM-KP43
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1
AUTHORS     Sato, T., Okuda, M., Koyama, S., Izawa, Y. and Kobayashi, T.
TITLE       Alkaline protease
JOURNAL     Patent: EP 1466970-A 2 13-OCT-2004;
            Kao Corporation (JP)
FEATURES
            Location/Qualifiers
            1..1923
            /organism="Bacillus sp. KSM-KP43"
            /mol_type="unassigned DNA"
            /db_xref="taxon:109322"
            1..1923
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAH61612.1"
            /db_xref="GI:54304587"
            /translation="MRKKKVPISLVLSAAIISTVALSNPDSAGANFDLDFKGIQTT
            :TDKGFSGKQGTGAAAFVSESENVKLPKGIQKLETVPANRKLIIQNPPIERTKO
            QLEKTKSLIDYIPDYAIVYEYGVKASSTIEHVESVPIYRIDPOLFTKAS
            ELVKAVALDPTKONKEVOLRGIQIOAFISNDVLYITAEKVMNDVARGVVKQDV
            AOSVGLYGGGQIVAVADPGLDGRBDSNHEA.PRGKTTALALGRTNANDNGHGT
            HVAQSVLGNSTKGMAPOMLVFQSTWDSGGELGGLPSNLQTLFSDVSAQARHIN
            SWGAANVGAATTDNRVDDIVRNDMTIIFPAQNEBPNGGTISAQTAANAIVGATE
            NMAPSGSYADININVAOFPSRPTDGR.IKPDVMA.PGTFLISARSLAPDSFPMNH
            DSKYATMGSTNAPPIVAGNVQLRHIFVNRGITPEPSLKLALAGAADIGLQYEN
            GNGKMGRTLDKSLNVA.YVNESLSLSTSOATYSPATACKPKILGVMSDASITTA
            SIVLVNDLIDIVITAPNGTQYVGNDFPSYNDNMDGRNENVEINAPQSGTTIIEVQA
            YNVPVGPQPSLAIVN"
            1..618
            619
            sig_peptide
            mat_peptide
            ORIGIN
            Query Match      100.0%; Score 1923; DB 6; Length 1923;
            Best Local Similarity 100.0%; Pred. No. 0;
            Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
            QY      1 ATGAGAAAGAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGAGATTTTGTGACT 60
            DB      1 ATGAGAAAGAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGAGATTTTGTGACT 60
            QY      61 GTTGCGTTAAGTAATCATCTGACAGGTGTGCAAGAAATTTGATCTGAATTTCAAGGA 120
            DB      61 GTTGCGTTAAGTAATCATCTGACAGGTGTGCAAGAAATTTGATCTGAATTTCAAGGA 120
            QY      121 ATTGACAGCAACATGATGTTTCTTCAAGGAGGAGGAGTGTGCTGCTGCT 180
            DB      121 ATTGACAGCAACATGATGTTTCTTCAAGGAGGAGGAGTGTGCTGCTGCTGCT 180
            QY      121 ATTGACAGCAACATGATGTTTCTTCAAGGAGGAGGAGTGTGCTGCTGCT 180
            DB      121 ATTGACAGCAACATGATGTTTCTTCAAGGAGGAGGAGTGTGCTGCTGCT 180
            QY      181 TTTCTGTGGAATCTGAAAATGTGAAATCCCAAAAGTTTGCAGAAAGAGCTTGAACA 240
            DB      181 TTTCTGTGGAATCTGAAAATGTGAAATCCCAAAAGTTTGCAGAAAGAGCTTGAACA 240
            QY      241 GTTCCGGCAATATATAATCTCATATTATCAATTCAATGACCAATTTTGAAGAAACA 300
            DB      241 GTTCCGGCAATATATAATCTCATATTATCAATTCAATGACCAATTTTGAAGAAACA 300
            QY      301 AAACACAGCTGAAAAACAGGGGCAAGATTTCTGACATACATACCTGATTTGCTTAC 360
            DB      301 AAACACAGCTGAAAAACAGGGGCAAGATTTCTGACATACATACCTGATTTGCTTAC 360
            QY      361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGAAACAGCATTGAGCAGTGAATCC 420

```

```

DB      361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGAAACAGCATTGAGCAGTGAATCC 420
QY      421 GTGAGGCTTATTTGCCGATATACAGAAATGATCCCGACTTTTACAAAAAGGGGATCA 480
DB      421 GTGAGGCTTATTTGCCGATATACAGAAATGATCCCGACTTTTACAAAAAGGGGATCA 480
QY      481 GAGCTGTAAACAGAGGCGCTTGATACAAACAGAAAAATTAAGGGGCAATTAGA 540
DB      481 GAGCTGTAAACAGAGGCGCTTGATACAAACAGAAAAATTAAGGGGCAATTAGA 540
QY      541 GGCATCGAACAAATCGCACATTCGCAATAGCAATGATGTGATATATTACCGCAAG 600
DB      541 GGCATCGAACAAATCGCACATTCGCAATAGCAATGATGTGATATATTACCGCAAG 600
QY      601 CTTGATATTAAGTATGATGATGATGTTGGCGGTGGAATTGTCAAGCGGATGCGTCA 660
DB      601 CTTGATATTAAGTATGATGATGATGTTGGCGGTGGAATTGTCAAGCGGATGCGTCA 660
QY      661 AGCAGCTACGGGTTGATGAGCAAGGACAGATGCTGACGGTTCGATACAGGGCTTGAT 720
DB      661 AGCAGCTACGGGTTGATGAGCAAGGACAGATGCTGACGGTTCGATACAGGGCTTGAT 720
QY      721 ACAGTTCGCAATGACAGTTGATGATGATGAAAGCTTCGCGGGAAAAATTACTGATTTAT 780
DB      721 ACAGTTCGCAATGACAGTTGATGATGATGAAAGCTTCGCGGGAAAAATTACTGATTTAT 780
QY      781 GCATTGGGACGACGAATTAATGCCAATGATACGAATGTCATGTCATGTCATGTCGCTGC 840
DB      781 GCATTGGGACGACGAATTAATGCCAATGATACGAATGTCATGTCATGTCATGTCGCTGC 840
QY      841 TCCGTTATTTGGAAACGGCTCCATTAATTAAGAAATGGCGCTCAGCGGAATCTAAGTTC 900
DB      841 TCCGTTATTTGGAAACGGCTCCATTAATTAAGAAATGGCGCTCAGCGGAATCTAAGTTC 900
QY      901 CAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB      901 CAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY      961 TTCAGCCAAAGCATACAGTGTGTCAGAAATTCATACAACTCTGCGGAGAGCAGTGT 1020
DB      961 TTCAGCCAAAGCATACAGTGTGTCAGAAATTCATACAACTCTGCGGAGAGCAGTGT 1020
QY      1021 AATGGGGCTTACCAACAATTCAGAAATGCGAATGCGAATGCGAATGCGAATGCGAATGCG 1080
DB      1021 AATGGGGCTTACCAACAATTCAGAAATGCGAATGCGAATGCGAATGCGAATGCGAATGCG 1080
QY      1081 ACGATCTTTTGTGCTGCTGCGGGAAATGAAGACGAAACGGCGGAACATCACTGACACAGGC 1140
DB      1081 ACGATCTTTTGTGCTGCTGCGGGAAATGAAGACGAAACGGCGGAACATCACTGACACAGGC 1140
QY      1141 ACGCTTAAATATGCAATTAACGTCGAGCTACGAAACCTCGCCCAAGCTTTGGGTC 1200
DB      1141 ACGCTTAAATATGCAATTAACGTCGAGCTACGAAACCTCGCCCAAGCTTTGGGTC 1200
QY      1201 TATGCGGAATATGCAACATGTCGACAGTTCTCTTCAAGTCGTCGTCGTCGTCGTCGTC 1260
DB      1201 TATGCGGAATATGCAACATGTCGACAGTTCTCTTCAAGTCGTCGTCGTCGTCGTCGTC 1260
QY      1261 CGGATCAAAACCGGATGTCATGTCGACCGGGAAAGTTTCAATCTATCAGCAAGATCTTCTCT 1320
DB      1261 CGGATCAAAACCGGATGTCATGTCGACCGGGAAAGTTTCAATCTATCAGCAAGATCTTCTCT 1320
QY      1321 GCAACGGATCTCTCTCTTCTGCGGGAACATGACAGTAAATATGATATGATGATGATGATG 1380
DB      1321 GCAACGGATCTCTCTCTTCTGCGGGAACATGACAGTAAATATGATATGATGATGATGATG 1380
QY      1381 TCCATGCTACACCGATGCTGTCGGAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1440
DB      1381 TCCATGCTACACCGATGCTGTCGGAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1440
QY      1441 AACAGAGGATCAACCAAGCTTCTCTATTTAAAGCGGCACTGATTCGCGGTGAGCT 1500
DB      1441 AACAGAGGATCAACCAAGCTTCTCTATTTAAAGCGGCACTGATTCGCGGTGAGCT 1500

```

QY	1501	GACATCGGCTTGGCTACCGCAACGGTAA	CAAAGATGGGAGCAGTGCATTGGATGAA	1566
Db	1501	GACATCGGCTTGGCTACCGCAACGGTAA	CAAAGATGGGAGCAGTGCATTGGATGAA	1566
QY	1561	TCCTCGAAGCTTGCCCTATGTGAACGAGTCAAGTCTCTATGCCACCAAGCCAAAACCGACG		1620
Db	1561	TCCTCGAAGCTTGCCCTATGTGAACGAGTCAAGTCTCTATGCCACCAAGCCAAAACCGACG		1620
QY	1621	TACTGTTACTGCTACTGCTCCGGCAAGCCTTGGAAATCTCCCTGGTATGGTCTGATGCC		1680
Db	1621	TACTGTTACTGCTACTGCTCCGGCAAGCCTTGGAAATCTCCCTGGTATGGTCTGATGCC		1680
QY	1681	CCTGGAGACACACTGCTCCGTAAACGCTTGCAATGATCTGGACCTTGCTCATACCGCT		1740
Db	1681	CCTGGAGACACACTGCTCCGTAAACGCTTGCAATGATCTGGACCTTGCTCATACCGCT		1740
QY	1741	CCAAATGGACACACTATGTATGGAATACTTACTCTTGCCCTATACAAATGATACCTGGAT		1800
Db	1741	CCAAATGGACACACTATGTATGGAATACTTACTCTTGCCCTATACAAATGATACCTGGAT		1800
QY	1801	GGCCGCAATACGTAGAAATGATATTTATTTATATGACACCAAGCGGACGATATCAATT		1866
Db	1801	GGCCGCAATACGTAGAAATGATATTTATTTATATGACACCAAGCGGACGATATCAATT		1866
QY	1861	GAGGTACAGGCTTATTAACGTACCGGTTGGACACAGACCTTCTGTTGGCAATTGTGAT		1920
Db	1861	GAGGTACAGGCTTATTAACGTACCGGTTGGACACAGACCTTCTGTTGGCAATTGTGAT		1920
QY	1921	TAA	1923	
Db	1921	TAA	1923	
RESULT 3				
LOCUS	CO889309	1923 bp	DNA	linear
DEFINITION	Sequence 2 from Patent EP1466962.			PAT 19-Oct-2004
ACCESSION	CO889309			
VERSION	CO889309.1	GI:54305176		
KEYWORDS				
SOURCE				
ORGANISM	Bacillus sp. KSM-KP43			
REFERENCE	Bacillus sp. KSM-KP43			
AUTHORS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
TITLE	Okuda, M., Sato, T., Takimura, Y., Sumitomo, N. and Kobayashi, T.			
JOURNAL	Alkaline protease Patent: EP 146962-A 2 13-Oct-2004; Kao Corporation (JP)			
FEATURES				
SOURCE	location/Qualifiers			
	1..1923			
	/organism="Bacillus sp. KSM-KP43"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:109322"			
	1..1923			
	/note="unamed protein product"			
	/codon_start=1			
	/transl_table=1			
	/protein_id="CAH1686.1"			
	/db_xref="GI:54305177"			
	1..1923			
	/translation="MRKKKKVFLSVLSAAAILSTVALSNPSAGARNFDLPFGIOIT			
	TDAGPSKGGOTGGAALFLVSEBNKLPKGIQKLETPYANNKLIHIOGPIIETKQ			
	OLEKTKAKITLDYIPDYAVIYVEBGDYKASSTIEHVSVEPYLPYRIDQLTKGAS			
	ELYKAVALDTRKQKKEVQLNGIEIDIAFPALSNDLVLTATPERKVMNDVARGIVKDV			
	AQSSYGLVGGQIVLVAADLTGIDTRGNSMHEAFRGKITLVYALGRNNANDNGHGT			
	HVAGSYVNGSITNGKAPQANLVFQSIIMDSGGIGGLPSMLQTLFSQAYSAGARITN			
	SNRGAIVAGATYDTSRNDVYRKNDMTILFAAGNGBNGGTISAPGTAKXVALTVGAT			
	NLWSPFSYADININHWAFQSSRGPTDGRIPKPYMAAGPTLISARSLAPDSFMANH			
	DSCTAYNGGSMATPIYAGNVADCLREHYVNRGRTITPRPSILKALALGAADIGLGP			
	GNQGWGVTITLDKSLNAYIVNESSLSLSTQATYSPTATACKPLKISLWSDAFASTTA			
	SVTLVNDLDTVITAPNGTQVYNDPFSYNDNWDGRNNVENFINAPQSGTYYTIEVOA			
	YNPVGGQFTPSLAIIVN"			
	1..618			

81g peptide

mat_peptide	619..1920	
ORIGIN	/product="unnamed"	
Query Match	100.0%; Score 1923; DB 6; Length 1923;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1923; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
QY	1 ATGAGAAAGAGAAAAAGGTGTTTTATCTGTTTTATCAGCTGCAGCGATTTGTGACT	60
DB	1 ATGAGAAAGAGAAAAAGGTGTTTTATCTGTTTTATCAGCTGCAGCGATTTGTGACT	60
QY	61 GTTGCGTTAAGTATCATCTGCAGGTGTGCAAGGAATTTGATCTGGAATTTCAAGA	120
DB	61 GTTGCGTTAAGTATCATCTGCAGGTGTGCAAGGAATTTGATCTGGAATTTCAAGA	120
QY	121 ATTCAACAACAACCTGATGCTAAAGGTTCTCAAGCAGGGCAGACTGTGCTGCT	180
DB	121 ATTCAACAACAACCTGATGCTAAAGGTTCTCAAGCAGGGCAGACTGTGCTGCT	180
QY	181 TTTCGTGTGGAAATCTGAAAATGTGAAATCTCCAAAAGTTTGCAAGAACTTGAAAC	240
DB	181 TTTCGTGTGGAAATCTGAAAATGTGAAATCTCCAAAAGTTTGCAAGAACTTGAAAC	240
QY	241 GTCCCGCAAAATATTAATCTCATATTATTCAAATTCAAATGCACAATTTTGAAGAA	300
DB	241 GTCCCGCAAAATATTAATCTCATATTATTCAAATTCAAATGCACAATTTTGAAGAA	300
QY	301 AAACGACAGCTGAGAAAAACAGGGCCAAAGATTCTGCATCATCTGATTAGCTTAC	360
DB	301 AAACGACAGCTGAGAAAAACAGGGCCAAAGATTCTGCATCATCTGATTAGCTTAC	360
QY	361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAACCATTTAGACAGTGAATCC	420
DB	361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAACCATTTAGACAGTGAATCC	420
QY	421 GTGAGCCTTATTTGCCGATATACAAATAGATCCCGCTTTTCACAAAAGGGCATCA	480
DB	421 GTGAGCCTTATTTGCCGATATACAAATAGATCCCGCTTTTCACAAAAGGGCATCA	480
QY	481 GAGCTGTAAAGACGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTAAGA	540
DB	481 GAGCTGTAAAGACGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTAAGA	540
QY	541 GGCATCGAACCAATGCGCACAATTGCGCAATGATGTGCTATATATTAACGGCAAG	600
DB	541 GGCATCGAACCAATGCGCACAATTGCGCAATGATGTGCTATATTAACGGCAAG	600
QY	601 CCTGATATTAAGTGAATGAATGATGTTCCGGTGGAAATTTGCAAAAGCGAATGTGCTAG	660
DB	601 CCTGATATTAAGTGAATGAATGATGTTCCGGTGGAAATTTGCAAAAGCGAATGTGCTAG	660
QY	661 AGCAGCTACGGGTGTATGACAAGACAGATCGTACGGTGTGCGATACAGGGCTTAT	720
DB	661 AGCAGCTACGGGTGTATGACAAGACAGATCGTACGGTGTGCGATACAGGGCTTAT	720
QY	721 ACAGGTGCGAATGACAGTTGCGATGATGACCTTTCCGGGAAAAATTACTGCATTATAT	780
DB	721 ACAGGTGCGAATGACAGTTGCGATGATGACCTTTCCGGGAAAAATTACTGCATTATAT	780
QY	781 GCATTGGGACGGAAGAAATGATGCAATGATGATGATGATGATGATGATGATGATGATG	840
DB	781 GCATTGGGACGGAAGAAATGATGCAATGATGATGATGATGATGATGATGATGATGATG	840
QY	841 TCCGTTATTAGAAAACGGCTCCATATATTAAGAAATGGGCTCAGGGGAATCTAGTCTTC	900
DB	841 TCCGTTATTAGAAAACGGCTCCATATATTAAGAAATGGGCTCAGGGGAATCTAGTCTTC	900
QY	901 CAATCTATCATGATACGGGTGGGGGACTTGAGAGACTTACTTGGAAATCTGCAAACTTAA	960
DB	901 CAATCTATCATGATACGGGTGGGGGACTTGAGAGACTTACTTGGAAATCTGCAAACTTAA	960
QY	961 TTCAGCAAGCATCAGTGTGTTGGGCAAGATTCATTCAAATCTCTGGGGAGCAGCAAGT	1020

Db 961 TTCAGCCAGCATACAGTGTGTCGACGAATTCATACAAATCTCTGGGAGAGCAGCGT 1020
Qy 1021 AATGGGGCTTACACACAGATTCACGAATATGTGATGATCTATGTGGCAAAAATGATATG 1080
Db 1021 AATGGGGCTTACACACAGATTCACGAATATGTGATGATCTATGTGGCAAAAATGATATG 1080
Qy 1081 ACAGTCTCTTTCGCTGCGGGGAATGAAGACCGGAGCGGGAACCATCAGTGCACAGCG 1140
Db 1081 ACAGTCTCTTTCGCTGCGGGGAATGAAGACCGGAGCGGGAACCATCAGTGCACAGCG 1140
Qy 1141 ACAGCTAAAAATGCAATTAACAGTCGAGCTACGGAACCTCCGCCCAAGCTTTGGGCT 1200
Db 1141 ACAGCTAAAAATGCAATTAACAGTCGAGCTACGGAACCTCCGCCCAAGCTTTGGGCT 1200
Qy 1201 TATGCGGACATATCAACCATGTGGCAAGTCTTCTTCACTGAGACCGAAGATGGA 1260
Db 1201 TATGCGGACATATCAACCATGTGGCAAGTCTTCTTCACTGAGACCGAAGATGGA 1260
Qy 1261 CGGATCAAAACCGGATGTCATGAGCAGCGGGAACGTTACATACAGAAAGATCTTCT 1320
Db 1261 CGGATCAAAACCGGATGTCATGAGCAGCGGGAACGTTACATACAGAAAGATCTTCT 1320
Qy 1321 GCAACCGGATTCCTCTTCTGGGCGAAACATGACAGTAATATGCAATAGGGTGAACG 1380
Db 1321 GCAACCGGATTCCTCTTCTGGGCGAAACATGACAGTAATATGCAATAGGGTGAACG 1380
Qy 1381 TCCATGCGCTACACCATTCCTGCTGGAACAGTGGCAAGCTTCTGAGCAATTTTGAAA 1440
Db 1381 TCCATGCGCTACACCATTCCTGCTGGAACAGTGGCAAGCTTCTGAGCAATTTTGAAA 1440
Qy 1441 AACAGAGGATGACACCAAGCCTTCTATTAAAGCGGCACTGATGCGGATGAGCT 1500
Db 1441 AACAGAGGATGACACCAAGCCTTCTATTAAAGCGGCACTGATGCGGATGAGCT 1500
Qy 1501 GACATGCGCTTGGCTACCGGACCGGTAAACAGAGTGGGAGCAGTGCATTTGATTA 1560
Db 1501 GACATGCGCTTGGCTACCGGACCGGTAAACAGAGTGGGAGCAGTGCATTTGATTA 1560
Qy 1561 TCCCTGAACGTTGCTATGTGAACAGTGCATGTTCTTATCCACGCGCAAAAAGCGAG 1620
Db 1561 TCCCTGAACGTTGCTATGTGAACAGTGCATGTTCTTATCCACGCGCAAAAAGCGAG 1620
Qy 1621 TACTGCTTACTGCTACCTGCGGCAAGCCTTGAATAATCCCTGATGATGCTGATGCC 1680
Db 1621 TACTGCTTACTGCTACCTGCGGCAAGCCTTGAATAATCCCTGATGATGCTGATGCC 1680
Qy 1681 CCTGGAGACCAACTGCTTCCGTAACGCTTGTCAATGATCTGGAACCTTGTCAATCCGCT 1740
Db 1681 CCTGGAGACCAACTGCTTCCGTAACGCTTGTCAATGATCTGGAACCTTGTCAATCCGCT 1740
Qy 1741 CCAATGCGACACATGATGAGAAATGACTTTACTTCCCATACATGATTAATCGGAT 1800
Db 1741 CCAATGCGACACATGATGAGAAATGACTTTACTTCCCATACATGATTAATCGGAT 1800
Qy 1801 GGCCTGAATTAAGTGAATAATGATTTAATGACCAAAAGCGGAGCTTATTAAT 1860
Db 1801 GGCCTGAATTAAGTGAATAATGATTTAATGACCAAAAGCGGAGCTTATTAAT 1860
Qy 1861 GAGGTACAGGCTTATACGTAACGCTTGGACCAAGACCTTCTGTTGCAATTTGAT 1920
Db 1861 GAGGTACAGGCTTATACGTAACGCTTGGACCAAGACCTTCTGTTGCAATTTGAT 1920
Qy 1921 TAA 1923
Db 1921 TAA 1923

RESULT 4
LOCUS AR368117 1923 bp DNA 1linear PAT 12-SEP-2003
DEFINITION Sequence 5 from patent US 6376227.
ACCESSION AR368117

VERSION AR368117.1 GI:34601778
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa, M., Okuda, M., Saeiki, K., Kubota, H., Hitomi, J., Kageyama, Y.,
Shikata, S. and Nomura, M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 5 23-APR-2002;
Kao Corporation; Tokyo;
JPX;
FEATURES
source Location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1923; DB 6; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAAAGAGAAAGAGGTTTATCTGTTTATCAGCTGACGATTTTGTGACT 60
Db 1 ATGAGAAAGAGAAAGAGGTTTATCTGTTTATCAGCTGACGATTTTGTGACT 60
Qy 61 GTTGCCTTAAGTATCATCTGACGCTGTCAGAGGAATTTGATCTGATTTCAAGGA 120
Db 61 GTTGCCTTAAGTATCATCTGACGCTGTCAGAGGAATTTGATCTGATTTCAAGGA 120
Qy 121 ATTGAGACAACTGATGCTAAAGGTTCTCCAAACAGGGGACGCTGCTGCT 180
Db 121 ATTGAGACAACTGATGCTAAAGGTTCTCCAAACAGGGGACGCTGCTGCT 180
Qy 181 TTTCTGCTGAATCTGAAATATGTAATCTCCCAAAAGTTTGGAGAAGACTTGAACA 240
Db 181 TTTCTGCTGAATCTGAAATATGTAATCTCCCAAAAGTTTGGAGAAGACTTGAACA 240
Qy 241 GTTCCGGCAATTAATCACTCCATTTATTCATATGACCAATTTTGAAGAAACA 300
Db 241 GTTCCGGCAATTAATCACTCCATTTATTCATATGACCAATTTTGAAGAAACA 300
Qy 241 GTTCCGGCAATTAATCACTCCATTTATTCATATGACCAATTTTGAAGAAACA 300
Db 241 GTTCCGGCAATTAATCACTCCATTTATTCATATGACCAATTTTGAAGAAACA 300
Qy 301 AAAACGACCTGGAAGAAACAGGGGCAAGATTTCTGACTACATCTGATTAATGCTTAC 360
Db 301 AAAACGACCTGGAAGAAACAGGGGCAAGATTTCTGACTACATCTGATTAATGCTTAC 360
Qy 361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGAAGACGTGAATCC 420
Db 361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGAAGACGTGAATCC 420
Qy 421 GTTGAACCTTATTTGCGATATACAGATGATGCCAGCTTTTCAAAAAGGGGATCA 480
Db 421 GTTGAACCTTATTTGCGATATACAGATGATGCCAGCTTTTCAAAAAGGGGATCA 480
Qy 481 GAGCTTGTAAAGCAGTGGCGCTTGAATCAAAAGCAGAAATTAAGAGGTGCAATTAAGA 540
Db 481 GAGCTTGTAAAGCAGTGGCGCTTGAATCAAAAGCAGAAATTAAGAGGTGCAATTAAGA 540
Qy 541 GGCATCGAAACAATGCGCAATTTCCCAATTAAGATGATGCTATATTAACGGAAG 600
Db 541 GGCATCGAAACAATGCGCAATTTCCCAATTAAGATGATGCTATATTAACGGAAG 600
Qy 541 GGCATCGAAACAATGCGCAATTTCCCAATTAAGATGATGCTATATTAACGGAAG 600
Db 541 GGCATCGAAACAATGCGCAATTTCCCAATTAAGATGATGCTATATTAACGGAAG 600
Qy 601 CCTGAGTAAAGTGAATGATGTTGGCGCTGAAATTTGCAAAAGCGATGTGCTCAG 660
Db 601 CCTGAGTAAAGTGAATGATGTTGGCGCTGAAATTTGCAAAAGCGATGTGCTCAG 660
Qy 661 AGCAGCTACGGGTGATGAGCAAGAGCAGATGTAAGCGGTTCCGATCAAGGGCTTGA 720
Db 661 AGCAGCTACGGGTGATGAGCAAGAGCAGATGTAAGCGGTTCCGATCAAGGGCTTGA 720
Qy 721 ACAGGTGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACAGGTGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGAT 780

Db 541 GGATCGAACAAATCGCAATTCGCAATAGCAATATATGTCATATATATACGGCAAG 600
Qy 601 CCGAGATTAAGGTGATGAATGATTTGGCCGTGGAAATTTGTCAAGCGAATGTGGCTAG 660
Db 601 CCGAGATTAAGGTGATGAATGATTTGGCCGTGGAAATTTGTCAAGCGAATGTGGCTAG 660
Qy 661 AGCAATTAAGGTGATGAATGATTTGGCCGTGGAAATTTGTCAAGCGAATGTGGCTAG 720
Db 661 AGCAATTAAGGTGATGAATGATTTGGCCGTGGAAATTTGTCAAGCGAATGTGGCTAG 720
Qy 721 ACAGGTGCAATGACAGTTGATGATGAAGCCTTCCGCGGAAATTTACTGCAATTAAT 780
Db 721 ACAGGTGCAATGACAGTTGATGATGAAGCCTTCCGCGGAAATTTACTGCAATTAAT 780
Qy 781 GCAATGGGACGACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 GCAATGGGACGACGAATTAATGCAATGATGATGATGATGATGATGATGATGATGATG 840
Qy 841 TCGGATTAAGGAAACGGCTCACTAATAAGAAATGGCGCTGAGGGAAATTTAGTCTTC 900
Db 841 TCGGATTAAGGAAACGGCTCACTAATAAGAAATGGCGCTGAGGGAAATTTAGTCTTC 900
Qy 901 CAATCTATCATGATAGCGGTGGGGAATTTGAGAGTACTTCTGCAATCTGCAAACTTA 960
Db 901 CAATCTATCATGATAGCGGTGGGGAATTTGAGAGTACTTCTGCAATCTGCAAACTTA 960
Qy 961 TTGAGCGACGATACAGTGTGTCGCAATTAATCAAACTCTGCGGAGACAGCATG 1020
Db 961 TTGAGCGACGATACAGTGTGTCGCAATTAATCAAACTCTGCGGAGACAGCATG 1020
Qy 1021 AATGGGGCTTACACAGATTCGAGAAATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 AATGGGGCTTACACAGATTCGAGAAATGATGATGATGATGATGATGATGATGATGATG 1080
Qy 1081 ACGATCTTTTCCGTCCGGGAAATGAAAGACCGAAGCGGAAACATCAATGACACAGC 1140
Db 1081 ACGATCTTTTCCGTCCGGGAAATGAAAGACCGAAGCGGAAACATCAATGACACAGC 1140
Qy 1141 ACGATCTTTTCCGTCCGGGAAATGAAAGACCGAAGCGGAAACATCAATGACACAGC 1200
Db 1141 ACGATCTTTTCCGTCCGGGAAATGAAAGACCGAAGCGGAAACATCAATGACACAGC 1200
Qy 1201 TATGCGGACAAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 TATGCGGACAAATCAATCAATGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1261 CGGATCTTTTCCGTCCGGGAAATGAAAGACCGAAGCGGAAACATCAATGACACAGC 1320
Db 1261 CGGATCTTTTCCGTCCGGGAAATGAAAGACCGAAGCGGAAACATCAATGACACAGC 1320
Qy 1321 GCAATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db 1321 GCAATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Qy 1381 TCCATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1381 TCCATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Qy 1441 AACAGAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 1441 AACAGAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy 1501 GACATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db 1501 GACATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Qy 1561 TCCATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db 1561 TCCATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Qy 1621 TACTCGTTTACTGCTAGCTGCGGCAAGCTTTGAAATATCTCTGATGATGATGATGATG 1680
Db 1621 TACTCGTTTACTGCTAGCTGCGGCAAGCTTTGAAATATCTCTGATGATGATGATGATG 1680

Db 1621 TACTCGTTTACTGCTAGCTGCGGCAAGCTTTGAAATATCTCTGATGATGATGATGATG 1680
Qy 1681 CCGCGACGACAACTGCTTCCGTAACGCTTGTGATGATGATGATGATGATGATGATGATG 1740
Db 1681 CCGCGACGACAACTGCTTCCGTAACGCTTGTGATGATGATGATGATGATGATGATGATG 1740
Qy 1741 CCAATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
Db 1741 CCAATGGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
Qy 1801 GCGCGAATTAAGGTGATGAATGATTTTATGATGATGATGATGATGATGATGATGATG 1860
Db 1801 GCGCGAATTAAGGTGATGAATGATTTTATGATGATGATGATGATGATGATGATGATG 1860
Qy 1861 GAGGTACAGGCTTATTAAGTACGATGATGATGATGATGATGATGATGATGATGATG 1920
Db 1861 GAGGTACAGGCTTATTAAGTACGATGATGATGATGATGATGATGATGATGATGATG 1920
Qy 1921 TAA 1923
Db 1921 TAA 1923
RESULT 6
AR368118
LOCUS AR368118 1923 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 7 from patent US 6376227.
ACCESSION AR368118
VERSION AR368118.1 GI:34601779
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 7 23-APR-2002;
Kao Corporation; Tokyo;
JPK;
FEATURES
source location/Qualifiers
1..1923
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.8%; Score 1919.8; DB 6; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGCAAGATTTGTGACT 60
Db 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGCAAGATTTGTGACT 60
Qy 61 GTTGGCTTAAGTATCATCTGACAGGTGTGCAAGAAATTTGATGTGATTTCAAGGA 120
Db 61 GTTGGCTTAAGTATCATCTGACAGGTGTGCAAGAAATTTGATGTGATTTCAAGGA 120
Qy 121 ATTGAGACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Db 121 ATTGAGACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 181 TTTCTGGTGAATCTGAAATATGAAATCTCCCAAAAGTTTGCAGAAAGATTTGAAACA 240
Db 181 TTTCTGGTGAATCTGAAATATGAAATCTCCCAAAAGTTTGCAGAAAGATTTGAAACA 240
Qy 241 GTCCGGCAATTAATTAATCTCCATATATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 GTCCGGCAATTAATTAATCTCCATATATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 AAAAGACAGCTGGAAGAAACAGGGGCAAGATTTCTGACATGATGATGATGATGATGATG 360
Db 301 AAAAGACAGCTGGAAGAAACAGGGGCAAGATTTCTGACATGATGATGATGATGATGATG 360

QY 361 ATTGTCAGATATGAGGGCGATGTTAAAGTCAGCAACAGGACCATGAGACGCGTGAATCC 420
DB 361 ATTGTCAGATATGAGGGCGATGTTAAAGTCAGCAACAGGACCATGAGACGCGTGAATCC 420
QY 421 GTGAGACCTTTATTTGCGATATACAGAAATAGATCCCAGCTTTTCACAAAAGGGGCAATCA 480
DB 421 GTGAGACCTTTATTTGCGATATACAGAAATAGATCCCAGCTTTTCACAAAAGGGGCAATCA 480
QY 481 GAGCTGTAAAGCAATGGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCATTTAA 540
DB 481 GAGCTGTAAAGCAATGGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCATTTAA 540
QY 541 GGCATGCAACAAATCGCACATTCGCAATTAAGCATATGCTGCTATATATTAACGGCAAG 600
DB 541 GGCATGCAACAAATCGCACATTCGCAATTAAGCATATGCTGCTATATATTAACGGCAAG 600
QY 601 CCTGATATAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCTGATATAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGTACGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 AGCAGTACGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACAGGTGCGAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCGAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTTGGGACGACGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTTGGGACGACGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATATTAGAAACGGCTCCAATAAAGAAATGGCGCTCAGGCGCAATTAATGATGATGAT 900
DB 841 TCCGATATTAGAAACGGCTCCAATAAAGAAATGGCGCTCAGGCGCAATTAATGATGATGAT 900
QY 901 CAATCTATCATGATATGACGGTGGGAGCTTGGAGGACTTCTTCGAATCTGCAAACTTAA 960
DB 901 CAATCTATCATGATATGACGGTGGGAGCTTGGAGGACTTCTTCGAATCTGCAAACTTAA 960
QY 961 TTCACGCAAGCATACAGTCTGCTGCGCAGAAATTCATCAAACTCTCTGGGAGCAGCAGTG 1020
DB 961 TTCACGCAAGCATACAGTCTGCTGCGCAGAAATTCATCAAACTCTCTGGGAGCAGCAGTG 1020
QY 1021 AATGGGGCTTACACAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGGCTTACACAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCTTTTCTGCTGCGGGAATGAAGACCGAACGGCGGAACCATGATGATGATGATGATGAT 1140
DB 1081 ACGATCTTTTCTGCTGCGGGAATGAAGACCGAACGGCGGAACCATGATGATGATGATGATGAT 1140
QY 1141 ACAGCTAAATATGCAATACAGTCCGAGCTACGGAACCTCTCCGCCAAGCTTTGGGCTCT 1200
DB 1141 ACAGCTAAATATGCAATACAGTCCGAGCTACGGAACCTCTCCGCCAAGCTTTGGGCTCT 1200
QY 1201 TATGCGGACAAATATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 TATGCGGACAAATATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CGGATTAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 CGGATTAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GCACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 GCACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 TCCATGCGTACACCGATGCTGCTGCGGAACGTCGACAGCTTCTGAGCAATTTTGTGAAA 1440
DB 1381 TCCATGCGTACACCGATGCTGCTGCGGAACGTCGACAGCTTCTGAGCAATTTTGTGAAA 1440

QY 1441 AACAGAGCATCACCAAAAGCTTCTCTATTTAAAGGGGACATGATCCCGGTGACACT 1500
DB 1441 AACAGAGCATCACCAAAAGCTTCTCTATTTAAAGGGGACATGATCCCGGTGACACT 1500
QY 1501 GACATCGGCTTTGGCTACCCGGAACGGTAAACGAAGATGGGACGAGTACATTTGATTTAA 1560
DB 1501 GACATCGGCTTTGGCTACCCGGAACGGTAAACGAAGATGGGACGAGTACATTTGATTTAA 1560
QY 1561 TCCCTGAACGTTGGCTTATGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1561 TCCCTGAACGTTGGCTTATGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 TACTGTTTACTGCTACTGCGGCAAGGCTTTGAAATCTCCGTGATGATGATGATGATGATGATGAT 1680
DB 1621 TACTGTTTACTGCTACTGCGGCAAGGCTTTGAAATCTCCGTGATGATGATGATGATGATGATGAT 1680
QY 1681 CCGTGGACCAACATGCTTCGCTTACAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 CCGTGGACCAACATGCTTCGCTTACAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 CCAATGGCACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 CCAATGGCACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GGCCGCAATTAACGTAGAAATGATATTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 GGCCGCAATTAACGTAGAAATGATATTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 GAGGTACAGGCTTATTAACGATCGGTTGACACAGACCTTCTGTTGGCAATTTGAT 1920
DB 1861 GAGGTACAGGCTTATTAACGATCGGTTGACACAGACCTTCTGTTGGCAATTTGAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 7
ARS62466
LOCUS ARS62466 1923 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 7 from patent US 6759228.
ACCESSION ARS62466
VERSION ARS62466.1 GI:53976507
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1923)
AUTHORS Takaiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
Shikata,S. and Nomura,M.
TITLE Alkaline protease
JOURNALS Patent: US 6759228-A 7 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;
FEATURES
source location/Qualifiers
1..1923 /organism="unknown"
/mol_type="genomic DNA"

Query Match 99.8%; Score 1919.8; DB 6; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCACTGACGAGATTTTGTGACT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCACTGACGAGATTTTGTGACT 60
QY 61 GTTGCGTTAAGTATCATCATGACAGGTGAGCAAGAAATTTGATCTGATTTCAAGGA 120
DB 61 GTTGCGTTAAGTATCATCATGACAGGTGAGCAAGAAATTTGATCTGATTTCAAGGA 120
QY 121 ATTACAGCAACACTGATGCTAAAGGTTTCTCCAGCAGGGGCGACACTGCTGCTGCT 180

source
1. .1923
/organism="Bacillus sp. KSM-9865"
/mol_type="genomic DNA"
/strain="KSM-9865"
/db_xref="taxon:192495"
1. .1923
/codon_start=1
/transl_table=11
/product="protease"
/db_xref="GI:34392387"
/translation="MRKKKKVPLSVLSAAAILSTVALSNPSAGANPDLDPKGIQTI
TDKGFSGKSGOTGAALFVSESNVKKPKGLQKLETPVANNKHLIVQFNPILSTKQ
OLEKTKALIDYIPDYAVIVEYGVKSAISTLEDESYPEPLIYRIDLPOLPFTGAS
ELVKVALDPTNKNKEVOLRGIEOLAOFAFSTNDVLYITAKPEPKYNNDAVAGIVKADY
AASSYGLYQGGIVAVADTGLDPTGRDSSMHEAPRKITALYALGETNNANDTNGHT
HVAAGVLNGSTNKGA-PQANLVFOSIMSGGLGGLPSNLDITLFSQASARLHTN
SWGAVNAGYTTDSRNVDYVRKNMTTILFAAGNBPNGGTLSABGTAKNAITVATE
NLRPSFGSYADNINHVAFSSRGPEDGRIKPDVMAFGTFILSASSILAPDSFWANH
DSKYAMGGTSMATPIVAGNVAQLRHFVKNRGITPKPSILKALILAGADIGLGYPN
GNQGGKRVTLDSKLSINVAAYNESLSLSTGKATYSPTATKGRKLKISLWSDAPASTTA
STTVNDLDELVTATPNGTQYVGNDFTSPTNNMDEGRNVENVFINAPSGTYTIEVOA
INVPGPQTFSLAIVN"

ORIGIN

Query Match 98.2%; Score 1887.8; DB 1; Length 1923;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1901; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACT 60
DB 1 ATGAGAAAGAGAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACT 60
QY 61 GTTGGTTAAGTAATCCATCTGCAGGTGTCAGAGAAATTTTGATTTGCAAGGA 120
DB 61 GTTGGTTAAGTAATCCATCTGCAGGTGTCAGAGAAATTTTGATTTGCAAGGA 120
QY 121 ATTGACAGCAACAATGATCTAAAGGTTCTCCAGCAGGGGCAAGCTGGTCTGTCT 180
DB 121 ATTGACAGCAACAATGATCTAAAGGTTCTCCAGCAGGGGCAAGCTGGTCTGTCT 180
QY 181 TTTCTGTGTGAATCTGAAAAATGTGAACCTCCAAAAAGGTTTGCAGAGAGCTTGAAA 240
DB 181 TTTCTGTGTGAATCTGAAAAATGTGAACCTCCAAAAAGGTTTGCAGAGAGCTTGAAA 240
QY 241 GTCCCGGGCAATAATTAATCTCAATTAATCCAAATTCAGTGCATATTTAGAGAA 300
DB 241 GTCCCGGGCAATAATTAATCTCAATTAATCCAAATTCAGTGCATATTTAGAGAA 300
QY 301 AAACAGCAGCTGGAAGAAAAAGGGGCAAGATTTCTGACTACATACCTGATATGCTTAC 360
DB 301 AAACAGCAGCTGGAAGAAAAAGGGGCAAGATTTCTGACTACATACCTGATATGCTTAC 360
QY 361 ATTGTGAGATAGAGGGCGATTTAAGTCAGCAACAAGCACCATTGAGCAGTGAATCC 420
DB 361 ATTGTGAGATAGAGGGCGATTTAAGTCAGCAACAAGCACCATTGAGCAGTGAATCC 420
QY 421 GTGAGCGCTTAATTTGCCGATATACAGAAATAGATCCCGACGTTTTCACAAAAGGGGCA 480
DB 421 GTGAGCGCTTAATTTGCCGATATACAGAAATAGATCCCGACGTTTTCACAAAAGGGGCA 480
QY 481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCAATTAAGA 540
DB 481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAATTCAGAAAAATTAAGAGGTGCAATTAAGA 540
QY 541 GGCATCGAACAATTCGCAATTTGCGAACAAGCAGATGATGCTATATATTAACGCAAG 600
DB 541 GGCATCGAACAATTCGCAATTTGCGAACAAGCAGATGATGCTATATATTAACGCAAG 600
QY 601 CCTGAGTAATAGAGATGAATGATGCTGCGGTGGAATTTGCAAGCGGATGTGGCGAG 660
DB 601 CCTGAGTAATAGAGATGAATGATGCTGCGGATTTGCAAGCGGATGTGGCGAG 660

QY 661 AGCAGCTACGGGTGTATGACAAGAGACAGATCGTAGGGGTTCGGATACAGGGCTGAT 720
DB 661 AGCAGCTACGGGTGTATGACAAGAGACAGATCGTAGGGGTTCGGATACAGGGCTGAT 720
QY 721 ACAGGTGCAATGACAGTGTGATGATGAAGCTTCCGCGGAAAAATTAAGTATATAT 780
DB 721 ACAGGTGCAATGACAGTGTGATGATGAAGCTTCCGCGGAAAAATTAAGTATATAT 780
QY 781 GCATTGGGACGAGCAATTAATGCCCAATGATACGAATGATCATGTGATGCTGCGC 840
DB 781 GCATTGGGACGAGCAATTAATGCCCAATGATACGAATGATCATGTGATGCTGCGC 840
QY 841 TCCGTAATTAAGAAAAAGGCTCCACTTAATAAGAAATGGCGCTCAGGCGCAATCTAGCTTC 900
DB 841 TCCGTAATTAAGAAAAAGGCTCCACTTAATAAGAAATGGCGCTCAGGCGCAATCTAGCTTC 900
QY 901 CAATCTATCATGATAGCGGTGGGAGCTTGAAGGACTTACCTTGCAATCTGCAACCTTTA 960
DB 901 CAATCTATCATGATAGCGGTGGGAGCTTGAAGGACTTACCTTGCAATCTGCAACCTTTA 960
QY 961 TTCAGCCAGCATACAGTGTGCTGTCGCAAAATTCATACAACTCCTGGGAGCAGCAGTG 1020
DB 961 TTCAGCCAGCATACAGTGTGCTGTCGCAAAATTCATACAACTCCTGGGAGCAGCAGTG 1020
QY 1021 AATGGGGCTTACACAACAGATTCAGAAATGTGATGCTATGTGGCGCAAAATGATATG 1080
DB 1021 AATGGGGCTTACACAACAGATTCAGAAATGTGATGCTATGTGGCGCAAAATGATATG 1080
QY 1081 ACGATCCTTTTCGCTGCGGGAAATGAAGACCGAACCGGCAACATCAGTGCAACAGGC 1140
DB 1081 ACGATCCTTTTCGCTGCGGGAAATGAAGACCGAACCGGCAACATCAGTGCAACAGGC 1140
QY 1141 ACAGTAAATATGCAATTAACAGTGTGAGACTACGAAAACTCCGCCCAAGCTTTGGGCT 1200
DB 1141 ACAGTAAATATGCAATTAACAGTGTGAGACTACGAAAACTCCGCCCAAGCTTTGGGCT 1200
QY 1201 TATGGGCACAATTAATCAACATAGTGGCAAGTTCCTTCAAGTGGACCGACAAAGATGGA 1260
DB 1201 TATGGGCACAATTAATCAACATAGTGGCAAGTTCCTTCAAGTGGACCGACAAAGATGGA 1260
QY 1261 CGGATCAACCGGATGTGATGACCGGGAACTTCAATATCATGACAAAGATCTTCTCT 1320
DB 1261 CGGATCAACCGGATGTGATGACCGGGAACTTCAATATCATGACAAAGATCTTCTCT 1320
QY 1321 GCACCGGATTCCTCTCTTGTGGGCAACATGACATTAATGCAATACATGCTGGGAA 1380
DB 1321 GCACCGGATTCCTCTCTTGTGGGCAACATGACATTAATGCAATACATGCTGGGAA 1380
QY 1381 TCCATGGCTACACGATCGTGTGTAAGACGTGGCAAGCTTGTGAGCAATTTGTGAAA 1440
DB 1381 TCCATGGCTACACGATCGTGTGTAAGACGTGGCAAGCTTGTGAGCAATTTGTGAAA 1440
QY 1441 AACAGGCAATCACACCAAAAGCTTCTTAATTAAGCGGCACTGATTCGGGTGAGCT 1500
DB 1441 AACAGGCAATCACACCAAAAGCTTCTTAATTAAGCGGCACTGATTCGGGTGAGCT 1500
QY 1501 GACATCGGCTTGTGGCTACCGGAACGTGAACCAAGAGTGGGACGAGTGAACATTTGATAAA 1560
DB 1501 GACATCGGCTTGTGGCTACCGGAACGTGAACCAAGAGTGGGACGAGTGAACATTTGATAAA 1560
QY 1561 TCCCTGAACGTTGTGCTATGTGAACGATTCAGATTCCTATTCACACAGCAAAAAAGGAG 1620
DB 1561 TCCCTGAACGTTGTGCTATGTGAACGATTCAGATTCCTATTCACACAGCAAAAAAGGAG 1620
QY 1621 TACTGTTTACATGCTACCTGCGGCAAGCTTTGAAAAATCTCCCTGTATGCTGATGCC 1680
DB 1621 TACTGTTTACATGCTACCTGCGGCAAGCTTTGAAAAATCTCCCTGTATGCTGATGCC 1680
QY 1681 CTTGCGGACCAACTGCTTCCCTGTAACGCTTGTCAATGATGCACTTGTCAATTAACGCT 1740
DB 1681 CTTGCGGACCAACTGCTTCCCTGTAACGCTTGTCAATGATGCACTTGTCAATTAACGCT 1740
QY 1741 CCAAAATGGCACACAGTATGTAGAAATGACTTAACTTCGCGCATACATGATTAACGAGAT 1800

|||||
Db 1741 CCAATGGCACACATATGTTGGAAATGACTTACTTCGCGCATACATATTAATCGGAT 1800
Qy 1801 GCGCGCATATACGTAGAAAATGTAATTTATATGACCAACCAAGCGGAGTATACAAAT 1860
Db 1801 GCGCGCATATACGTAGAAAATGTAATTTATATGCGCCACCAAGCGGAGTATACAAAT 1860
Qy 1861 GAGTACAGGCTTATATACGTACCGGTGACCAAGACCTTCCTGCTTGGAATTTGAAAT 1920
Db 1861 GAGTACAGGCTTATATACGTACCGGTGACCAAGACCTTCCTGCTTGGAATTTGAAAT 1920
Qy 1921 TAA 1923
Db 1921 TAA 1923
RESULT 9
AB046403 1920 bp DNA linear BCT 10-MAY-2002
LOCUS Bacillus sp. 9860 PROA gene for protease, complete cds.
DEFINITION AB046403
ACCESSION AB046403.2 GI:20521152
KEYWORDS
SOURCE Bacillus sp. 9860
ORGANISM Bacillus sp. 9860
REFERENCE 1
AUTHORS Saeki,K., Okuda,M., Hatada,Y., Kobayashi,T., Ito,S., Takami,H. and
TITLE Novel oxidatively stable subtilisin-like serine proteases from
JOURNAL alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
PUBLISHED Biochem. Biophys. Res. Commun. 279 (2), 313-319 (2000)
REFERENCES 2 (bases 1 to 1920)
AUTHORS Saeki,K.
JOURNAL Direct Submission
TITLE Submitted (20-JUL-2000) Katuhisa Saeki, Kao corporation,
JOURNAL Biological Science Laboratory, Ichikaimachi Akabane 2606, Haga,
TELEPHONE Tel:81-285-68-7400, Japan (E-mail:1387185@kasetanet.kao.co.jp,
FAX Tel:81-285-68-7403)
COMMENT On May 9, 2002 this sequence version replaced gi:12381938.
FEATURES
source location/Qualifiers
1..1920
/organism="Bacillus sp. 9860"
/mol_type="genomic DNA"
/strain="9860"
/db_xref="taxon:133778"
1..1920
/gene="PROA"
1..1920
/gene="PROA"
/protein_id="BAB21266.2"
/cds_start=1
/cds_end=1920
/protein_id="BAB21266.2"
/db_xref="GI:20521153"
/translation="MRKKVFLVLSAAALSTVALNPSSAGDARTFDLFGKIQTT
DVSFSGKQRTGAALFLVSENVKLLKLETPANKKLIYVFNPILEETOK
LETTGAKLIDYIPVAYIVIEGSDVOSKRSIEHVSVEPVIKRYIDLPOLTKAST
LYKALALDPRKNKKVQLRGIERIAQYVANSVNHVITTAPEYKVMNDVARGIYKADVA
GSSYGLYGGQIYVAVDTGLDGRNDSNHEAFRGKITLVALAGRTNNANDNGHGH
VAGSVNGATNKGAPQANLVFQSIIDSSGGIGLPNMQTLFQSAFAGARIHNS
WGAANVAGATTSRVNDVVRKNDKTLIFAGNERNGGTSIPGAKNAITVGAEN
LSPFSGYADININVAQFSRGPETKGRIPKDVMAEGTILSARSLADSSFMAND
SKYAVGSGMATPIVAGNVAOLRHFVKNRGITTPKPSILKALTAGADVIGYVNG
NGMGKRVTLDEKLANVAVNESLSTSGAATTPPTATKPKLISLWMDAPASTAS
VTLVNDLDTITTPNCTRYVGNDSAPFNNDGRNNVNFINSQGTITTEVAY
NPFVGPQNSLAIWN"

ORIGIN
Query Match 76.8%; Score 1477.2; DB 1; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;

Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
Qy 10 AAAAAAAGGTGTTTATCTGTTTATCACTGACGACGATTTTGTGACTGTGCTTA 69
Db 7 AAAAAAAGGTGTTTATCTGTTTATCACTGACGACGATTTTGTGACTGTGCTTA 66
Qy 70 AGTAATCCATCTGACAGGTGTGCAAGAAATTTTGAATTCGATTTCAAGAAATTCAGACA 129
Db 67 AACAAATCCCTGCTGCTGTGATGCAAGAACTTTGAATTCGATTTTAAAGAAATTCAGACA 126
Qy 130 ACAATGATGCTTAAGTCTTCTCAAGCAGGCGACAGCTGCTCTGCTTCTGCTG 189
Db 127 ACACCGAATGTCTAGTGTCTTCTCAAGCAGGCGACAGCTGCTCTGCTGCTG 186
Qy 190 GAATCTGAAAATGTGAACTCCCAAAAGGTTTGCAAGAAAGCTTGAACAGTCCCGCA 249
Db 187 GAGTCTGAAAATGTGAACTCTTAAAGGATTTCTAAAGAACTTGAACAGTCCCGCA 246
Qy 250 AATAATTAATCTCAATTTATCCAAATTCATGACCAATTTTGAAGAAACAAACAGCAG 309
Db 247 AATAATTAATCTCAATTTATGCAATTCATGACCAATTTTGAAGAAACAAACAGCAG 306
Qy 310 CTGAAAAAACAAGGCGCAAGATTCGACTACATCTGATTAATGCTTACATTTGTCGAG 369
Db 307 CTGAGACAACTGAGACCAAGATTCGACTACATCTGATTAATGCTTACATTTGTCGAG 366
Qy 370 TATGAGGCGGATTTAATGTCAGCAACAGCAACCATTTGACGACGTGAATCCGTGAGCCT 429
Db 367 TATGAGGCGGATTTAATGTCAGCAACAGCAACCATTTGACGACGTGAATCCGTGAGCCT 426
Qy 430 TATTTCCGATATACGAATATGATCCCGACTTTTCAAAAAGGCGCATACAGCTTGTGA 489
Db 427 TACTTGCCCAAAATACAAATTAATGATCCCGACTTTTCAAAAAGGCGCATACAGCTTGTGA 486
Qy 490 AAAAGCAGTGGCGCTTATACAAAGCAGAAATTAAGAGGTGCAATTAAGAGGATCGAA 549
Db 487 AAGGCTTGGCGCTTATACAAAGCAGAAATTAAGAGGTGCAATTAAGAGGATCGAA 546
Qy 550 CAAATGCACAATTCGCAATTAAGCAATGATGCTATATTAACGCAAAAGCTGATAT 609
Db 547 GAAATGCTCAGTACGTACCAAGCAATGATGCTATATTAACGCAAAAGCTGATAT 606
Qy 610 AAGGTATGAATGATTTGCGGTGCAATTTGCAAAAGCGAGTGTGCTCAGACACTAC 669
Db 607 AAGGTATGAATGATTTGCGGTGCAATTTGCAAAAGCGAGTGTGCTCAGACACTAC 666
Qy 670 GGGTGTATGACAAGGACAGATTCGAGCGGTGCGCATACAGGCTGTATCAGGTCCG 729
Db 667 GGTGTGTATGACAAGGACAGATTTGCGAGTTGCGCATACAGGCTGTATCAGGTCCG 726
Qy 730 AATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
Db 727 AATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
Qy 790 CGAGCAATTAATGCAATGATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 849
Db 787 CGAGCAATTAATGCAATGATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 846
Qy 850 GGAAGCGGCTCACTAATAAGAAATGAGGCTCAGGCGGAATCTGTGCTCAATCTATC 909
Db 847 GGAAGCGGCTCACTAATAAGAAATGAGGCTCAGGCGGAATCTGTGCTCAATCTATC 906
Qy 910 ATGATATGCGGTGAGGAGCTTGAGAGACTTCTGCAATCTGCAACCTTATTCAGCCAA 969
Db 907 ATGATATGCGGTGAGGAGCTTGAGAGACTTCTGCAATCTGCAACCTTATTCAGCCAA 966
Qy 970 GCATACAGTCTGCTGCTGCAAGATTTATACAACTCTCTGGGAGCAGCAGTGAATGGGCT 1029
Db 967 GCATACAGTCTGCTGCTGCAAGATTTATACAACTCTCTGGGAGCAGCAGTGAATGGGCT 1026
Qy 1030 TACACAAAGATTCACAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1089
Db 1027 TACACAAAGATTCACAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1086

QY 1090 TTGGCTCCGGGGAATGAGGACCGGAAACGGCGGAACCATGATGTCACACAGGCAAGCTTAA 1149
DB 1087 TTGGCGGCTGGGATATGAAAGGCGGAAACGGGCGTATCATGTCACCTGGTACGGCTTAA 1146
QY 1150 AATGCAATTAACAGTCCGAGCTACGGAACCTCCGGCCAAAGCTTTGGTCTTATGCGGAC 1209
DB 1147 AACGCATTAACAGTCCGCGCAACCGAAACCTCCGTCGAAGCTTGGTCTTATGCAAGT 1206
QY 1210 AATATCAACCATGTGGCAAGTTCTCTTCACTGACCGAACAAGATGACGGAATCAA 1269
DB 1207 AATATTAACCAAGTTCACAGTCTCTCCGCGGCCCAACAAGATGGCGAATCAAG 1266
QY 1270 CCGGATGTCATGGCACCGGGAACGTCATATCTATGACGAAGATCTTCTTGACACCGGAT 1329
DB 1267 CCGATGTCATGGCGCGGGAACATATTTATTCAGAAAGATCTTCTTGACACCGGAT 1326
QY 1330 TCTCTCTTGGGCGGAACCATGACAGTAATATGATACATGGGTGGAAGCTTCATGCT 1389
DB 1327 TCTCTCTTGGGCGGAATCATGACAGCAATATGCTATATGGGTGGAAGCTTCATGCA 1386
QY 1390 AACACGATGTTGCTGGAACCTGGGCAAGCTTCTGAGCAATTTGTGAAAAACAGAGC 1449
DB 1387 AACCGATTTGTGCGGGGAATGTGGCAAGCTCCGAGCAATTTGTGAAAAATAGAGGA 1446
QY 1450 ATCAACCAAGAGCTTCTATTTAAAGGCGACTGATTGGCGGTCAGCTGACATCGGC 1509
DB 1447 ATCACTCTTAAAGCTTCTCTATTTGAAGCAAGCTTTGATGAGGTGCTGATGTTGA 1506
QY 1510 CTGGCTACCCGGAACCGTAACCAAGATGGGGAAGAGTGAATGATTAATCCCTGAC 1569
DB 1507 TTGGGTTATCCGAACGGAACCAAGATGGGCGGAGTGAACCTGGAATTAATCGTTGAC 1566
QY 1570 GTTGCCTATGTAACGAGTCAGTCTCTATTCACAGGCAAAAAGCGATCTGCTT 1629
DB 1567 GTTGCCTATGTAACGAGTCAGTCTCTATTCACAGGCAAAAAGCGATCTGCTT 1626
QY 1630 ACTGCTACTGCGGCAAGCTTTGAAAATCTCCCTGTAATGGTGTGATGCCCTGAGAC 1689
DB 1627 ACTGCAACGCGCGGCAAGCTTTGAAAATCTCCCTGTAATGGTGTGATGCCCTGAGAC 1686
QY 1690 AACACCTGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTATTAACGCTTCAATGAC 1749
DB 1687 ACTACTGCTTCTGTAAACCTGTGTCAATGATTTGATTTGTATTAACGACCAACGGA 1746
QY 1750 AACAGTATGTAAGAAATGATTTTATGATGACCAACGCGGACGTAATGATGATGAT 1809
DB 1747 AACAGTATGTAAGAAATGATTTTATGATGACCAACGCGGATGATGATGATGATGAT 1806
QY 1810 AACGTAGAAATGATTTTATGATGACCAACGCGGACGTAATGATGATGATGATGATGAT 1869
DB 1807 AACGTAGAAATGATTTTATGATGACCAACGCGGATGATGATGATGATGATGATGAT 1866
QY 1870 GCTTATTAACGATCCGCTTGGACCAAGACTTCTGCTTGGCAATTTGAAATTA 1923
DB 1867 GCATATTAATGTGCGGTTGGACCAACAACTTCTGCTTGGCAATTTGAAATTA 1920

RESULT 10
AR368116
LOCUS AR368116 1920 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6376227.
ACCESSION AR368116
VERSION AR368116.1 GI:34601777
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1920)
AUTHORS Takeiwa,M., Okuda,M., Saeki,K., Kubota,H., Hitomi,J., Kageyama,Y.,
TITLE Alkaline protease
JOURNAL Patent: US 6376227-A 3 23-APR-2002;

Kao Corporation; Tokyo;
JPK;
FEATURES
source Location/Qualifiers
1..1920
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 76.8%; Score 1477.2; DB 6; Length 1920;
Best Local Similarity 85.7%; Pred.No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 10 AAGAAAAGGTTTATCTGTTTATCACTGACGAGATTTGTGACCTGTCGTTA 69
DB 7 AAGAAAGGTTTATCTGTTTATCACTGACGAGATTTGTGACCTGTCGTTA 66
QY 70 AGTATCCATCTGACAGTGTGCAAGAAATTTTATCTGATTCGATTTCAAGAAATTCAGACA 129
DB 67 AACATCCCTCGGCTGATGATGCAAGACCTTTGATCTGATTTTAAAGAAATTCAGACA 126
QY 130 AACATGATCTAAAGTTTCTCCAAAGGAGGAGACTGAGTGTCTGCTTTCTGCTG 189
DB 127 AACACCGATGTCAAGTGTCTCCAAACAGGCAACAGAGTGGCTGATTTCTGCTG 186
QY 190 GAATCTGAATATGGAATCTCCAAAGGTTTGAAGAAAGCTTGAACATGCTCCGCA 249
DB 187 GAGTCTGAATATGGAATCTTAAAGGATTTGTAAGAAATTTGAACGTAACCGCA 246
QY 250 AATATTAATCTCAATATTTATCAATGACCAATTTTGAAGAAACAAACAGCAG 309
DB 247 AATATTAATCTCAATATTTATGTCATATTCATGATGATGATGATGATGATGATGATGAT 306
QY 310 CTGGAAGAAACAGGAGGCAAGATTTCTGATCACTATCTGATTTATGCTTACATTTGCTGAG 369
DB 307 CTAGAGCAACTGAGAGAAAGATTTCTGATCACTATCTGATTTATGATATATTTGCTGAG 366
QY 370 TATGAGGCGATGTTTAAAGTCAAGCAAGCACTTGAACGATGGAATCCGTGAGCCT 429
DB 367 TATGAGGCGATGTTTAAAGTCAAGCAAGCACTTGAACGATGGAATCCGTGAGCCT 426
QY 430 TATTTGCGATATACAGAAATGATCCCAAGCTTTTCAAAAGGAGGATGAGACTGTA 489
DB 427 TACTTCCGAAATACAAATATGATCCCAAGCTTTTCAAAAGGAGGATGAGACTGTA 486
QY 490 AAACAGTGGGCTTGTATCAAGACAGAAATTAAGAGTGCATTTAAAGGATGAA 549
DB 487 AAACAGTGGGCTTGTATCAAGACAGAAATTAAGAGTGCATTTAAAGGATGAA 546
QY 550 CAATTCGACAAATTCGCAATTAAGATGATGCTATATTAACGGAAGCCGATAT 609
DB 547 GAATTCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 610 AAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
DB 607 AAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 670 GGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
DB 667 GGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
QY 730 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
DB 727 AATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 790 CGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
DB 787 CGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 850 GGAAGGCTCTCACTAATTAAGAAATGAGGCTTCAAGGCAATTAATGATGATGATGATGATGATGATGATGATGAT 909
DB 847 GGAAGGCTCTCACTAATTAAGAAATGAGGCTTCAAGGCAATTAATGATGATGATGATGATGATGATGATGATGAT 906
QY 910 ATGATATGCGGTGGGGAATTTGAGAGACTTCTGATGATGCAATCTTATTCAGCCAA 969

Db 907 ATGATAGACAGTGGTGGCTTGAGGCTTGCTTCATCTGCAAACTTATTACGCA 966
Qy 970 GCATACAGTCTGTGTGCAAAATTCATACAACTCTGGGAGACAGACGATGGGCT 1029
Db 967 GCATACAGTCTGTGTGCAAAATTCATACAACTCTGGGAGACAGACGATGGGCT 1026
Qy 1030 TACACACAGATTCACAAATGTGTGATGATGTGTGCGCAAAATGTATGATGATGCTT 1089
Db 1027 TACACACAGATTCACAAATGTGTGATGATGTGTGCGCAAAATGTATGATGATGCTT 1086
Qy 1090 TTGCTGCGCGGAAATGAAGACCGAACGCGGAAACCATCAGTGCACAGGACAGCTTAA 1149
Db 1087 TTGCGCGCTGGGATGAAGAGCGGACCGGCGATCATCAGTGCACCTGGATGGCTTAA 1146
Qy 1150 AATGCATATACGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1209
Db 1147 AACGCCATTAACGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1206
Qy 1210 AATATCAACATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1269
Db 1207 AATATTAACATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1266
Qy 1270 CCGGATGTCAATGCAACCGGAGAGTTCATATACAGCAAGATCTTCTTGTGACCGGAT 1329
Db 1267 CCGGATGTCAATGCAACCGGAGAGTTCATATACAGCAAGATCTTCTTGTGACCGGAT 1326
Qy 1330 TCTCTCTTGTGGGCAACATGACATTAATATGATCATCATGTGGTGAAGCTTCAGGCT 1389
Db 1327 TCTCTCTTGTGGGCAACATGACATTAATATGATCATCATGTGGTGAAGCTTCAGGCT 1386
Qy 1390 ACACCGATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1449
Db 1387 ACACCGATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1446
Qy 1450 ATCACAACCAAGCTTCTCTATTAATAAGCGGCACTGATTCGCGTGCAGCTGCATCGGC 1509
Db 1447 ATCACAACCAAGCTTCTCTATTAATAAGCGGCACTGATTCGCGTGCAGCTGCATCGGC 1506
Qy 1510 CTGGCTACCGCAACCGTAAACCAAGATGGGAGAGTGCATGTGATTAATCTCGTGAAC 1569
Db 1507 CTGGCTACCGCAACCGTAAACCAAGATGGGAGAGTGCATGTGATTAATCTCGTGAAC 1566
Qy 1570 GTTGGCTATGTGAACAGTCCAGTCTCTATTCACAGCGCAAAAGGAGTACTGCTT 1629
Db 1567 GTTGGCTATGTGAACAGTCCAGTCTCTATTCACAGCGCAAAAGGAGTACTGCTT 1626
Qy 1630 ACTGCTACTGCGCGCAAGCTTGAATAATCTCCCTGGTATGTCTGATGCGCGTGCAGC 1689
Db 1627 ACTGCTACTGCGCGCAAGCTTGAATAATCTCCCTGGTATGTCTGATGCGCGTGCAGC 1686
Qy 1690 ACAACTGCTTCCGTAAACGCTTGTCAATGATGTGACCTTGTCAATACCGCTCCAAATGGC 1749
Db 1687 ACAACTGCTTCCGTAAACGCTTGTCAATGATGTGACCTTGTCAATACCGCTCCAAATGGC 1746
Qy 1750 ACAAGATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1809
Db 1747 ACAAGATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 1806
Qy 1810 AACGTAGAAATATATTTATTAATGACACCAAGCGGAGCTATTAATGATGATGAC 1869
Db 1807 AACGTAGAAATATATTTATTAATGACACCAAGCGGAGCTATTAATGATGATGAC 1866
Qy 1870 GCTTATTAAGTACCGGTGAGACCAAGCTTGTGTGCAATGTGATTA 1923
Db 1867 GCTTATTAAGTACCGGTGAGACCAAGCTTGTGTGCAATGTGATTA 1920

RESULT 11
LOCUS AR562464 1920 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6759228.
ACCESSION AR562464

VERSION AR562464.1 GI:53976505
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1920)
AUTHORS Takaiwa, M., Okuda, M., Saeki, K., Kubota, H., Hltomi, J., Kageyama, Y.,
Shikata, S. and Nomura, M.
TITLE Alkaline protease
JOURNAL Patent: US 6759228-A 3 06-JUL-2004;
Kao Corporation; Tokyo;
JPK;
FEATURES
source location/Qualifiers
1..1920
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 76.8%; Score 1477.2; DB 6; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
Qy 10 AAGAAAAGGTGTTTATCTGTTTATCAGTGCAGCCGATTTGTGACCTGTGCGTTA 69
Db 7 AAGAAAAGGTGTTTATCTGTTTATCAGTGCAGCCGATTTGTGACCTGTGCGTTA 66
Qy 70 AGTAATCATCTGACAGGTGTGCAAGAAATTTGATCTGATTTCAAGAAATTCAGACA 129
Db 67 AACAAATCCCTCGGCTGTGTGATGCAAGCACTTTGATCTGATTTAAAGAAATTCAGACA 126
Qy 130 ACACCGATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 189
Db 127 ACACCGATGTGTGAGTACGAAATCTCGCCCAAGCTTGGCTTATGCGGAC 186
Qy 190 GAATCTGAATAATGTGAACCTCCCAAGGTTTGCAGAGAGCTTGAACAGTCCGCGCA 249
Db 187 GAATCTGAATAATGTGAACCTCCCAAGGTTTGCAGAGAGCTTGAACAGTCCGCGCA 246
Qy 250 AATATTAATCTCCATATTAATTCATTAATGACCAATTTTGAAGAAACCAAAACAGCAG 309
Db 247 AATATTAATCTCCATATTAATTCATTAATGACCAATTTTGAAGAAACCAAAACAGCAG 306
Qy 310 CTGGAAAACCAAGGCGCAAGATTTCTGCACTATCATCTGATTAATGATGCTTGAAG 369
Db 307 CTGGAAAACCAAGGCGCAAGATTTCTGCACTATCATCTGATTAATGATGCTTGAAG 366
Qy 370 TATGAGGCGATGTGAATCAGCAACCAAGCTTGAAGCACTGAGAAATCCGTGAGCCT 429
Db 367 TATGAGGCGATGTGAATCAGCAACCAAGCTTGAAGCACTGAGAAATCCGTGAGCCT 426
Qy 430 TATTTGCGGATTAACGAATTAATGATCCAGCTTTTCAAAAAGGCGCATCAGACTTGA 489
Db 427 TATTTGCGGATTAACGAATTAATGATCCAGCTTTTCAAAAAGGCGCATCAGACTTGA 486
Qy 490 AAAAGCGTGGCGCTTATTAACGAAGCAAGAAATTAAGAGTGCATTAAGAGGATCGAA 549
Db 487 AAAAGCGTGGCGCTTATTAACGAAGCAAGAAATTAAGAGTGCATTAAGAGGATCGAA 546
Qy 550 CAAATGCAACATTTGCAATTAAGCAATGATGTGCTATTAATTAAGGCAAGCTGATAT 609
Db 547 CAAATGCAACATTTGCAATTAAGCAATGATGTGCTATTAATTAAGGCAAGCTGATAT 606
Qy 610 AAGGTGATGAATGATTTGCGGTGAAATTTCAAGCGGATGTGCTCAGAGCAGCTTAC 669
Db 607 AAGGTGATGAATGATTTGCGGTGAAATTTCAAGCGGATGTGCTCAGAGCAGCTTAC 666
Qy 670 GGGTGTATGACCAAGCAAGATCTGAGGCTTCCGATTAACAGGCTTATTAAGAGTCCG 729
Db 667 GGGTGTATGACCAAGCAAGATCTGAGGCTTCCGATTAACAGGCTTATTAAGAGTCCG 726
Qy 730 AATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
Db 727 AATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786

QY 790 CGGACGAATTAATGCAATGATAGCAATGTCATGTAACGATGTCGCTGCTCCGATTTA 849
 DB 787 CGGACGAATTAATGCAATGATAGCAATGTCATGTAACGATGTCGCTGCTCCGATTTA 846
 QY 850 GGAACCGCTCCATATTAAGGAATGCGCTCAGCGCAATAGTCTTCCATCTATC 909
 DB 847 GGAATAGCGGCAACGAATTAAGGAATGCGCTCAGCGCAATAGTCTTCCATCTATC 906
 QY 910 ATGGAATAGCGGCTGCGGCACTTGGAGATACCTTGCATCTGCAACCTTATTCACCA 969
 DB 907 ATGGAATAGCGGCTGCGGCACTTGGAGATACCTTGCATCTGCAACCTTATTCACCA 966
 QY 970 GCATACAGTCTGTCGCAAGATTCATCAAACTCCGCGGAGACAGATGGAATGGGCT 1029
 DB 967 GCATACAGTCTGTCGCAAGATTCATCAAACTCCGCGGAGACAGATGGAATGGGCT 1026
 QY 1030 TACACAACGATTCAGAAATGTCATGATGTCGCAAAATGATATGACATCTT 1089
 DB 1027 TACACAACGATTCAGAAATGTCATGATGTCGCAAAATGATATGACATCTT 1086
 QY 1090 TTGCTGCGGCAATGAGAACGAAACGCGGCAACATCATGTCACCAAGCTATA 1149
 DB 1087 TTGCTGCGGCAATGAGAACGAAACGCGGCAACATCATGTCACCAAGCTATA 1146
 QY 1150 AATGCAATTAACAGTCCGAGCTACGAAACCTCCGCGCAAGCTTTGGTCTTATGCGAC 1209
 DB 1147 AATGCAATTAACAGTCCGAGCTACGAAACCTCCGCGCAAGCTTTGGTCTTATGCGAC 1206
 QY 1210 AATTAACCATGTCGACAGTTCTTCAAGTGAACGCAAAAGATGGAACGATCAAA 1269
 DB 1207 AATTAACCATGTCGACAGTTCTTCAAGTGAACGCAAAAGATGGAACGATCAAA 1266
 QY 1270 CCGGATGTCATGCAACCGGGAACGTCATCATATGCAAGATCTTCTTTCACCGAT 1329
 DB 1267 CCGGATGTCATGCAACCGGGAACGTCATCATATGCAAGATCTTCTTTCACCGAT 1326
 QY 1330 TCCCTCTTCTGCGGCAACCATGACAGTAAATATGACATATGAGTGAACGTCATGCT 1389
 DB 1327 TCCCTCTTCTGCGGCAACCATGACAGTAAATATGACATATGAGTGAACGTCATGCT 1386
 QY 1390 ACAACGATGTCGTCGGAACGTCGACAGCTTCTGAGCATTTTGTGAAAAACAGAGC 1449
 DB 1387 ACAACGATGTCGTCGGAACGTCGACAGCTTCTGAGCATTTTGTGAAAAACAGAGC 1446
 QY 1450 ATCAACCAAGCCCTTCTTATTAAGGCGCATGATGTCGCGTGCAGCTACATCGC 1509
 DB 1447 ATCAACCAAGCCCTTCTTATTAAGGCGCATGATGTCGCGTGCAGCTACATCGC 1506
 QY 1510 CTGTCATACCGCAACGTCATCAAGATGCGGACGATGACATGATTAATCCCTGAC 1569
 DB 1507 TTGGGTTATCCGAACGGAACCAAGATGCGGACGATGACATGATTAATCCCTGAC 1566
 QY 1570 GTTGCCTATGTAACGAGTCGATCTCTATCCACCAAGCAAAACGATCTGCTT 1629
 DB 1567 GTTGCCTATGTAACGAGTCGATCTCTATCCACCAAGCAAAACGATCTGCTT 1626
 QY 1630 ACTGCTACGCGCGGAAGCTTTGAAATCTCCCTGATGATGCTGATGCTCCCTGAGC 1689
 DB 1627 ACTGCTACGCGCGGAAGCTTTGAAATCTCCCTGATGATGCTGATGCTCCCTGAGC 1686
 QY 1690 ACAACGCTTCCGTAACGCTTGTCAATGTCGATCTGACCTTGTCAATTCGCAATGAC 1749
 DB 1687 ACTACGCTTCCGTAACGCTTGTCAATGTCGATCTGACCTTGTCAATTCGCAATGAC 1746
 QY 1750 ACAAGTATGTAAGGAATGATCTTATCCGCAATCAATGATTAATCTGGATGCGCGCAT 1809
 DB 1747 ACAAGTATGTAAGGAATGATCTTATCCGCAATCAATGATTAATCTGGATGCGCGCAT 1806
 QY 1810 AACGTAAGTAATGATTAATTAATGACACCAAGCGGAGATGATTAATTAAGTACAG 1869
 DB 1807 AACGTAAGTAATGATTAATTAATGACACCAAGCGGAGATGATTAATTAAGTACAG 1866

QY 1870 GCTTATAACGTCACCGGTTGACACCAACGACTTCTCTGTTGGCAATTGTGAATTA 1923
 DB 1867 GCATTAATATGTCGCGTTGACACCAAACTTCTGTTGGCAATTGTGAATTA 1920
 RESULT 12
 AX839476
 LOCUS AX839476 1305 bp DNA linear PAT 15-DEC-2003
 DEFINITION Sequence 2 from Patent EP1347044.
 ACCESSION AX839476
 VERSION AX839476.1 GI:39922766
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus sp. KSM-KP43
 Bacillus sp. KSM-KP43
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE
 1 Okuda, M.K., Sato, T.K., Saito, K.K., Sumitomo, N.K., Izawa, Y.K., Sasaki, K.K., Kobayashi, T.K. and Nomura, M.K.
 TITLE Alkaline protease
 JOURNAL Patent: EP 1347044-A 2 24-SEP-2003;
 FEATURES
 source location/Qualifiers
 1..1305
 /organism="Bacillus sp. KSM-KP43"
 /mol_type="unassigned DNA"
 /db_xref="taxon:109322"
 <1..1305
 /note="unassigned protein product"
 /codon_start=1
 /transl_table=1
 /protein_id="CA85146.1"
 /db_xref="GI:39922767"
 /translation="NDVARGIVKADVAQSSYELYGQIVAVADTGLDTRDSSMHE
 AFRRKILALYALGRTNANDPTNGHTVAVSVLNGSTNGMAFQANLVQISLDSG
 GLGGLPNNLQTLFQAYAGARLHTNSMGAIVNGAYTDSRNDVYRKDMILLFPA
 GNEPNGTGISAPETAKALIVGATEMLRSPFSYADININVAOFSGRPTKGRIRP
 DVNAPGFTILSARSLAPDSSPMANHDSTKAVNGGTSMAPIVAGVAOLREHVKR
 GTPKPSLILKALILAGAADIGLCPYRNGQKRYTLDKSLNAVYVSSSLSTSQKT
 YSPFATNGEPLKISLVMSDAPASTTASVTLLVNDLVTAPNCTGVNDPFTSPYND
 WDRNNVENVFINAPDSGYTITVEQAVNVPGTFSLAVN"
 ORIGIN
 Query Match 67.9%; Score 1305; DB 6; Length 1305;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 619 AATGATGTCGCGCTGGAATTTGTCAAGCGGATGTCGTCAGACGATCGGCTGTAT 678
 DB 1 AATGATGTCGCGCTGGAATTTGTCAAGCGGATGTCGTCAGACGATCGGCTGTAT 60
 QY 679 GGAACGACGATCGTAGCGGCTTGCATACAGGCTTGAATACAGTCCGATGACAGT 738
 DB 61 GGAACGACGATCGTAGCGGCTTGCATACAGGCTTGAATACAGTCCGATGACAGT 120
 QY 739 TCGATGATGAAGCTTCCGCGGGAATAATTACGCAATTATATGCAATTGGGACGGAAT 798
 DB 121 TCGATGATGAAGCTTCCGCGGGAATAATTACGCAATTATATGCAATTGGGACGGAAT 180
 QY 799 AATGCCAATGATGAATGTCATGTCATGTCATGTCGCTCCGATTAAGAAACGCG 858
 DB 181 AATGCCAATGATGAATGTCATGTCATGTCATGTCGCTCCGATTAAGAAACGCG 240
 QY 859 TCCACTAATAAGGAATGCGGCTCAGGCGAATCTAGCTTTCATATTCATGATGATGC 918
 DB 241 TCCACTAATAAGGAATGCGGCTCAGGCGAATCTAGCTTTCATATTCATGATGATGC 300
 QY 919 GGTGGGGGACTTGGAGGACTACCTGGAATCTGCAACCTTATTCAGCCAGGATACAGT 978
 DB 301 GGTGGGGGACTTGGAGGACTACCTGGAATCTGCAACCTTATTCAGCCAGGATACAGT 360
 QY 979 GCTGGTCCAGAAATTCATCAAACTCTGCGGAGACAGACGATGATGGGCTTACACACA 1038
 DB 361 GCTGGTCCAGAAATTCATCAAACTCTGCGGAGACAGACGATGATGGGCTTACACACA 420


```

QY 912 GGATAGCGGTGGGGGACCTTGAGAGACCTCTGCAATCTGCAAAACCTTATTCAGCAAGC 971
  |||||
DB 1760 GGATATGTTGGAGGGCTGGAGAGACTACCTGCTTAATCTACAAACCTTATTCAGCAAGC 1819
QY 972 ATACAGTGTGTGGTGCAGAAATTCATCAAACTCTGGGGAGCAGCATGAAATGGGGCTTA 1031
  |||||
DB 1820 ATATAGTGTGGAGCGAGAAATTCATCAAACTCTGGGGAGCTCCAGTAAACGGTGCCTTA 1879
QY 1032 CCAACAGANTCCAGAAATGTGATGATCTATGCGCCAAAATGATATGACATCTCTTTT 1091
  |||||
DB 1880 TACGAGAGACTCTCGAAATGTGATGATTAATGGAATAAAATGATATGACATCTCTTTT 1939
QY 1092 CCGTGGCGGAAATGAGAGACCGAAGCGGGAACCATCAGTGCACGAGCAAGCTTAAAA 1151
  |||||
DB 1940 TCGCGCGGAAATGAGAGACCGAGTACCGGTACATCAGTGCACGAGCAAGCAAAAA 1999
QY 1152 TCGAATACAGTGCAGCTACCGAAGAAACCTCCGCGCAAGCTTTGGTCTTAATGCGGACAA 1211
  |||||
DB 2000 TCGAATACAGTGCAGCTACCGAAGAAACCTACGTCAGCTTGAATGCGGATTA 2059
QY 1212 TATCAACAGTGCAGCTCTCTCAAGTGCACGAGCAAGAGATGACGATCAAAC 1271
  |||||
DB 2060 TATCAACAGTGCAGCTCTCTCAAGTGCAGAGTCTCTAGAGATGACGATTAAGCC 2119
QY 1272 GGAATGCATGAGCACCGGGAACGTTCACTATCAGCAAGATCTTCTTGACCGGATTC 1331
  |||||
DB 2120 GGAATGCATGAGCACCGAGTACGATATCTCTCTGATATCATCTTATGCTCCAGATTC 2179
QY 1332 CTCCTTCTGGGGGAAACCATGACAGTAATATGCAATCATGGGTGGAACGTCATGCTTAC 1391
  |||||
DB 2180 CTCATCTCTGGGGGAAACCATGATATGTAATATGCTCTCATGGGTGATCTTATGCTTAC 2239
QY 1392 ACCGATCGTGTGGAAGGTGCAAGCTGTGTGCAATTTTGTGAAAACAGAGGCAT 1451
  |||||
DB 2240 TCCAAATGTAAGCAGTAATGTGCACAATTAAGGGGCAATTTGTGAAAATAGAGGGGT 2299
QY 1452 CACACCAAGCCTTCTCTATTAAGCGGCACTGATTTGCGGTGACGTGACATCGGCT 1511
  |||||
DB 2300 AACTCTTAAGCCTTCTCTTAAAGCTGCTTAAATGAGGTGCTGCGGATGTTGAGACT 2359
QY 1512 TGGCTAACCGGAACGTAACCAAGATGGGAGCAGTGCATTTGATTAATCCCTGAACGT 1571
  |||||
DB 2360 TGGCTTAACCAATGTGTAACCAAGATGGGGAAGATGATTAATCCCTGAATGT 2419
QY 1572 TGCCTATGTAAGCAGTGCAGTCTCTATCAACGAGCAAAAGAGGCTACCTGCTTAC 1631
  |||||
DB 2420 CGCATTTGTGTAAGTAAGAGAGCCCTTATCAACAGTCAAAAAGCAATATTCGTTAC 2479
QY 1632 TGTCTACGCGGCAAGCCTTTGAAAATCTCCCTGTATGCTGATGCCCCCTGCGAGCAC 1691
  |||||
DB 2480 GGCCTACAGCTGTGTAACCTTTAAAAATATCACTTGTGTGTCAGATGCAACGAGTAC 2539
QY 1692 AACTCTTCCGTAAGCCTTGTCAATGATTTGACCTTGTCAATTCGCTCAATGAGCAC 1751
  |||||
DB 2540 GACGGCATCACTAACCTTTAGTAATGATTTAGACTTAATCATCTGCACCAAAATGAGAC 2599
QY 1752 ACAAGTATGTAAGAAATGACTTTACTTGCATACATGATTAATGCGGAATGGCGCAATTA 1811
  |||||
DB 2600 TAAATACGTGCGAAATGACTTTTACAGCACCGTAATGATTAACATTTGGAGTGGAGAAACA 2659
QY 1812 CGTAGAAAAATGATTTATTAATGACCAACAAAGCGGAGGTATACATTTGAGAGTACAGGC 1871
  |||||
DB 2660 CGTAGAAAAATGATTTATTAATGATCTCTCAAAAGCGGAGGTATACAGTGGAAAGTCAAGC 2719
QY 1872 TTAATACGTAACCGGTGAGACCAAGACCTTCTGCTTGGCAATTTGGAATTA 1923
  |||||
DB 2720 TTAATACGTAACGTAAGTACCGCAAAACCTTTCTTAAAGATTTGATTAATTA 2771
  |||||

```

RESULT 14
 BD062155 3003 bp DNA linear PAT 27-Aug-2002
 LOCUS BD062155
 DEFINITION Nucleic acids encoding a polypeptide having protease activity.

```

ACCESSION BD062155
VERSION BD062155.1 GI:22607760
KEYWORDS JP 2001514529-A/39.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3003)
AUTHORS Sloma, A. and Christlanston, L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patient: JP 2001514529-A 39 11-SEP-2001;
NOVO NORDISK BIOTECH INC
COMMENT PN JP 2001514529-A/39
PD 11-SEP-2001
PF 09-JUN-1998 JP 1999503145
PI 12-JUN-1997 US 08/873479
PR ALAN SLOMA, LYNNIE CHRISTIANSON
PC C12N15/57, C12N15/75, C12N9/54, C12K14/00
CC Strandedness: Single;
CC Topology: Linear;
FEATURES
source Location/Qualifiers
1..3003
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 60.7%; Score 1166.4; DB 6; Length 3003;
Best Local Similarity 76.0%; Pred. No. 2,26-287;
Matches 1453; Conservative 0; Mismatches 456; Indels 3; Gaps 1;
QY 12 GAAAAAGGTGTTTATCTGTTTATCAGCTGCAGCAATTTGTCAGCTGTGCTTAAG 71
  |||||
DB 863 GAAAGGGGTTTTTTATCCGTTTATCAGTGTGCACATAATGCTTCTGTTGCTTTAAG 922
QY 72 TAATCCATCTCAGGTGTGCAAGAAATTTGATCTGATTTCCAAAGAAATTCAGACAA 131
  |||||
DB 923 CAGTCTTCTACTAATTTGGGGGCAACAAATTTGAAATTTGACTTTAAGGGGATGAGAACT 982
QY 132 AACTGATCTTAAAGGTTTCTCCAAAGCAGGGGCAAGCTGTGCTGCTTTTCTGTGGA 191
  |||||
DB 983 TACGCTAGAGAAAGCTGTCACCAAGCAAGAAACGGAAGGATCTTTTCTTGTAAA 1042
QY 192 ATCTGAAATGTGAAATCTCCAAAGGTTTGCAGAGAGAGCTTGAAACAGTCCCGGCAA 251
  |||||
DB 1043 CTCTGAAATGTGAAATCTCCAAAGAGATTTCAAAAGAAACTAGAAAGTATGTCAGCGGA 1102
QY 252 TAATTAACCTCATATTTATCCAAATTCATGAGCAATTTTAAAGAAACAAACAGCACT 311
  |||||
DB 1103 TAAAGCTATATATTCGTTCAATTTGACGAGCTTATTTTAAAGAAACGCAACTTCAACT 1162
QY 312 GGAATAAAACAGGGGCAAGATTTCTGACTACATACCTGATTTATGCTTACATTTGTCAGATA 371
  |||||
DB 1163 AGAGAAAGCGGAGCGAAGAAATTTCTGATTAATACATACAGATTAACCTTATATTTGCGATA 1222
QY 372 TGAAGGCGATTTTAATGTCAGCAACAGACCATTTGAGACGTGGAATCCGTGAGCCTTA 431
  |||||
DB 1223 TGAATGGGATGTAAAGCCGTAACTAAACGCAATTTGCGCATTTGGAATCGTTGAACATA 1282
QY 432 TTTGCCGATTAACGAATATGATCCCGAGCTTTTCAAAAAGGGGCACTCAGCTTGTAAA 491
  |||||
DB 1283 TTAACCTTTATATTAATTAATGACCGCAATATTTTTCAGAGAGACCTTCTGATTAATGAGA 1342
QY 492 AGCAGTGGCGCTTGATACAAAGCAGAAAAATTAAGAGGTGCAATTAAGAGGATCGAACA 551
  |||||
DB 1343 AACAGTACCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1402
QY 552 AATGCAATTTGCAATTAAGCAATGATGCTATATTAATTAATTAATTAATTAATTAATTAATTA 611
  |||||
DB 1403 AATGCCCCAATACCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1462
  |||||

```

OY	612	GGTGAATGAATGATGTTGGCGCTGGAAATGTCAAAGCGAATGGCTCAGAGCAGCTTACG	671
Db	1463	AGTTTGGAAATGACGCGGCCCGGGGCAATGTGTGAAMGACAACGTCGCAAAAATTAATCTTTGG	15222
OY	672	GTTGATATGAGCAAGAGCAGATCGTAGCGGTTGGCCGATACAGGGCTTGTATACAGTCCGCA	731
Db	1523	CTTATATATGACAAAGACAGATTTGTAGCAGTTGCTGATCTGGGCTTGTATACAGGAAGAA	15822
OY	732	TGACAGTTGATGATGAAAGCCTTCCGGGGAAAATTACTGCATTAATATGCAATTGGGACG	791
Db	1583	TGACAGTTGATGATGAAAGCAATTCGGGGTAAGTTAACCGACATATATGCACTGGGACG	16422
OY	792	GACCAATTAATGCCATGATGCAATGATGATGATGCAATGGTGGCTCCGATTAGG	851
Db	1643	AACGATATTAACCGCAATGATTCMAATGAGCAATGGAACCATGTTGCTGATCTGTATAGG	17022
OY	852	AAACGGCTCCCACTAATTAAGGAATGGGGGCTCAGGCGAATCTAGCTTCCATCTATCAT	911
Db	1703	AAATG--CTACAAATTAAGGGATGGACCCGCAACCAATCTAGCTTTCAATCTATAT	17592
OY	912	GGATAGCGGTGGGGGACTTGGAGCACTACTTCGATCTGCAAACTTATTCAGCAGC	971
Db	1760	GGATAGTGTGGAGGGCTGGGAGCACTACCTGCTAATTTACAAACATTAATTCAGTCAGC	18192
OY	972	ATPACAGTCTGGTGCAGAAATTCATACAACTCTCGGGAGCAGACGTGAATGGGGCTTA	10312
Db	1820	ATATAGTCTGGAGCGGAATTCATACAAATTCAGGGGGGCTCAAGTAACGGTGCTTA	18792
OY	1032	CACACAAATTCAGAAATGTGGAATGATATGTGGCAAAATGATATGACAGATCTTTT	10912
Db	1880	TACGACAACTCTCGAAATGTGATGATTAATGTGAAAAAATGATATGACAGATCTTTT	19392
OY	1092	CGCTGCCGGGAATGAAAGACCGAACGGCGGAACCATCAGTGCACCAAGGCAAGCTTAAAA	11512
Db	1940	TGCGGCCGGGAATGAGGGAACAGATAGCGGTATCAATGATGCACAGGAACAGCAAAAA	19992
OY	1152	TGCAATPACAGTCGGAGCTACGAAAACTCTCGGCCCAAGCTTTGGGTCTTAATGCGGCAA	12112
Db	2000	TGCGATPACAGTGGGGGACCAACCAAACTACGCTCAAGCTTCGATCTTAATGCGGATPA	20592
OY	1212	TATCAACATGTGGACAGTTCTCTCAACGTGACCGGCAAAAGATGGAACGATCAAAAC	12712
Db	2060	TATTAACCATGTGTCAATTTCTTCAAGAGTCTTACAGAGTGAACGTATTAAGCC	21192
OY	1272	GGATGTCAATGCAACCGGAGCGTTCAATCTATACAGCAAGATCTTCTTTCGACCGGATTC	13312
Db	2120	GGAGTCAATGCAACAGGTATGTAATTTCTCTGTGTAGATCATATTAAGCTCCAGATTC	21792
OY	1332	CTCCTTCTGGGCGAACATGACAGTAATGATATGATACATGAGGTGMAACGTCATGGCTAC	13912
Db	2180	CTCATTTCTGGGCAAAACATGATGATTAATATGCTCAATATGAGTGTACTTATAGGCTAC	22392
OY	1392	ACCGATGTTGTCTGGAACGTGGCACAGCTTCGTGAGCATTTTGTGAAAAACAAGGCAT	14512
Db	2240	TCCAAATGTATCAAGGTATATGTTGCAATTAAGGAGCATTTTGTGAAAAATAAGGGGT	22992
OY	1452	CACACAAAGCCTTCTCATTTAAAAAGGGGCACTGATTCGCGTGAAGCTGACATCCGCT	15112
Db	2300	AACTCTTACGCTTCTCTTTTAAAAAGCTGCTTATATGCAAGTGTGCGAGATTTGAACT	23592
OY	1512	TGGCTAACCGACGATTAACAAAGATGGGGAAGAGTGAATTTGTATTAATCCCTGAACGT	15712
Db	2360	TGGCTTTCCAATGTGATTAACAAAGATGGGGAAGAGTGAATTTGTATTAATCCCTGAACGT	24192
OY	1572	TGCTTAATGAAACAGTTCAGTTCTCTATCCACAGCCAAAAAGCAGATCTCTGTTTAC	16312
Db	2420	CGCAATTTGTATGAAAGAAAGACCCCTTATATCAACAAGCTCAAAAACGAATTTGTTTAC	24792
OY	1632	TGCTTAATGAAACAGTTCAGTTCTCTATCCACAGCCAAAAAGCAGATCTCTGTTTAC	16912
Db	2480	GGCTCAAGCTGTATTAACCTTTAAAAATATCACTGTTTGGTCAAGTGAACAGGTAGAC	25392
OY	1692	AAGTGTTCGATTAACGCTTGTCAATGATCTGGAACCTTGTCAATTAACGCTCCAAATGGAC	17512

Db	2540	GACGGCACTACTACTTATAGTAATGATTTAGACTTGTATCTACGCAACCAATGAGAC	2599
Qy	1752	ACAGATATATAGAAATGACTTTACTTGCGCCATACATATGATTAATCTGGATGGCCGCAATTA	1811
Db	2600	TAAATACGTCCGAATATGACTTTATACAGCACCGCTATGATATACATTTGGATGGCAGAAACA	2659
Qy	1812	CGTAGAAATATATTTATTTAAATGACCAAGAGCGGAGCGTATACATTTAGATGACAGCC	1871
Db	2660	COTGAAATATGTTTATTCATATGCTCTCCAAAGCGAAGCGTATACATGTCGAAGTGCAGGC	2719
Qy	1872	TTATACATACCGGTGGACCAAGACCTTCTCGTGGCAATTTGTAATTA	1923
Db	2720	TTACATATACAGTAAGTCCGCAACCTTTCTTTAGCGATGATACATTA	2771
RESULT 15			
E03808	E03808	2218 bp	DNA linear PAT 29-SEP-1997
LOCUS	E03808 DNA encoding alkaline protease Ya.		
ACCESSION	E03808 GI:2172022		
VERSION	JP 1992197182-A/1.		
KEYWORDS	Bacillus sp.		
SOURCE	Bacillus sp.		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
REFERENCE	Tobe S., Odera M. and Asai Y.		
AUTHORS	1 (bases 1 to 2218)		
TITLE	DNA CODING ALKALINE PROTEASE YA ENZYME AND PRODUCTION OF ALKALINE PROTEASE YA USING THE DNA		
JOURNAL	Patent: JP 1992197182-A 1 16-JUL-1992;		
COMMENT	LION CORP		
	OS	Bacillus sp.	
	PN	JP 1992197182-A/1	
	PD	16-JUL-1992	
	PF	28-NOV-1990 JP 1990327110	
	PI	TOBE SEIICHI, ODERA MOTOKYASU, ASAI YOSHIO	
	PC	C12N15/57, C11D3/386, C12N9/54, (C12N15/57, C12R1:07), (C12N9/54, C12R1:07);	
	CC	strandedness: Double;	
	CC	topology: linear;	
	CC	hypothetical: No;	
	CC	anti-sense: No;	
	CC	*source: strain=Y strain;	
	CC	*source: clone=pub8A;	
	FH	key	
	FH	location/Qualifiers	
	FT	CDS	
	FT	218..2125	
	FT	/product='precursor of alkaline protease Ya'	
	FT	sig_peptide	
	FT	218..823	
	FT	mat_peptide	
	FT	824..2122	
	FT	/product='Precursor of alkaline protease Ya'.	
FEATURES	Location/Qualifiers		
source	1..2218		
	/organism="Bacillus sp."		
	/mol_type="genomic DNA"		
	/db_xref="taxon:1409"		
ORIGIN			
	Query Match	48.7%;	Score 936.2; DB 6; Length 2218;
	Best Local Similarity	69.4%;	Pred. No. 2.1e-228;
	Matches 1335;	Conservative 0;	Mismatches 573; Indels 15; Gaps 4;
Qy	1	ATGAGAAAGAAAGAAAGCGCTTTTATCTGTTTATACACTGACGAGATTTGTGCACT	60
Db	218	ATGAGAGGGGAAAAAAGATGATGCTATACATGATGCTTCCGCTGCACTTAGCTCTCA	277
Qy	61	GTTACGTTAATGATATCATCTGCAGGTGGTGCAGAAATTTGATCTGATATTTCAAGAA	120
Db	278	GTTATGTTAGTACCACTACAGTGGGGCAG-----ATTTCAGAGATTTTAAATGCT	331
Qy	121	ATTGAGCAACAACTGATGCTAAAGGTTTCTCCAAAGAGGGGCGAGACTGATGCTGCTGCT	180

Db 332 GTGAAAGTTTGAATAATGCTA---CTTTGTTAAACGATTAAGTACGGGTGAGCATCC 388
 Qy 181 TTTCTGTGTTGGAATCTGAAAAATGTGAAATCCCAAAAGTTTGAGAGAAAGCTTGAAACA 240
 Db 389 TTTCTAGTAGAATACGAAAAATATTAATATTCCTAAAGGTATTCAAAAGAAAGCTAGAGACA 448
 Qy 241 GTCCGGGCAATTAATTAATCTCCATATATTCATTCATTCGATTCGAAATTTTGAAGAAACA 300
 Db 449 GTACAGAGAGATTAACAACTCTACATCTGTACAAATTTTACGACCAATTCAGAGAGAGAG 508
 Qy 301 AAACAGAGCTGGAAGAAAG 360
 Db 509 CCAAAAGAGTAGAGTCTAG 568
 Qy 361 ATTGTGAGATAGAGGCGATGTTAAGTACAGCAAGAGCAATTCAGAGAGAGAGAGAGAG 420
 Db 569 ATTTGTCAGTAT---AGTGTGCTACAAAATAATATAGTATTCATCTCTTGAAGAC 625
 Qy 421 GTGAGAGCTTATTTGGCCGATATACAGATAGATCCCGAGCTTTTCAAAAAGAGAGAGAG 480
 Db 626 GTACAAACCTTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 685
 Qy 481 GAGCTTGTAAAG 540
 Db 686 CAGCTTGTCAAGCGGTATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 745
 Qy 541 GGCAATCGAAG 600
 Db 746 GATTAGATGAGATGTTCAATATATCTGCAATTAATTAATTAATTAATTAATTAATTAAT 805
 Qy 601 CCTGAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 806 CCCGAGTATGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
 Qy 661 AGCAGCTACGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 Db 866 AACATTTACGAGTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 925
 Qy 721 ACAGTCCCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 Db 926 ACAGTCCCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
 Qy 781 GCATTTGAG 840
 Db 986 GCGTTAG 1045
 Qy 841 TCCGATTTAG 900
 Db 1046 TCTGATCTTGTGTAATG---CTTTAAATTAAGAGATGAGCTCCGCAAGCTTAATAGTCTTC 1102
 Qy 901 CAATCTATCATGATAG 960
 Db 1103 CAATCTATTAATGATAG 1162
 Qy 961 TTCAGCCAAAGATACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 Db 1163 TTTAGTCAAGCTTGAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1222
 Qy 1021 AATGGGCTTACAG 1080
 Db 1223 AATGAGAGCTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
 Qy 1081 AGGATCTTTTCCGTCGAG 1140
 Db 1283 AGGATCTTTTCCGTCGAG 1342
 Qy 1141 ACAGCTAAAGATGAG 1200
 Db 1343 ACAGGAG 1402
 Qy 1201 TATGCGAGAGATATACAG 1260
 Db 1403 ATAGCAGATTAACCAATCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1462

Qy 1261 CGGATCAAAACCGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 Db 1463 CGAATTTAAGCCGAG 1522
 Qy 1321 GCACCGAGATCTCTCTGAG 1380
 Db 1523 GCTTCAGACTCTTCTGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1582
 Qy 1381 TCCATGCTACACCGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Db 1583 TCCATGAG 1642
 Qy 1441 AACAGAGAGATCAACCAAGAGCTTCTATTAATTAATTAATTAATTAATTAATTAAT 1500
 Db 1643 AATGAGAGATTAATCTTAAGCTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 1702
 Qy 1501 GACATCGGCTTGTGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
 Db 1703 GATGTTGTTTAAAGATATCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1762
 Qy 1561 TCCCTGAACGTTGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 Db 1763 TCGTTAAATGATGAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1822
 Qy 1621 TACTGCTTACTGCTACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 Db 1823 TATTGCTTCAAG 1882
 Qy 1681 CCTGAG 1740
 Db 1883 CTTGAG 1942
 Qy 1741 CCAATGAG 1800
 Db 1943 CCGAATGAG 2002
 Qy 1801 GCGCGCAATTAACGAG 1860
 Db 2003 GGTGCAACAAATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2062
 Qy 1861 GAGGTACAGGCTTAAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
 Db 2063 GAGGTTCAAGGATTAATTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2122
 Qy 1921 TAA 1923
 Db 2123 TAA 2125

Search completed: April 8, 2006, 06:48:17
 Job time : 9427.85 secs

stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d) its ability to digest casein is not inhibited by oleic acid; (e) it has molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be used as enzymes in washing compositions for use in automatic dishwashers and for washing clothes. The stability to oxidizing agents allows the enzyme to be an effective component of washing compositions including bleaches. The present sequence represents an alkaline protease encoding CC DNA. (Updated on 20-MAR-2003 to correct DR field.)

XX Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;

Query Match 100.0%; Score 1923; DB 2; Length 1923;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGAAAGAGAAAGGTTTATCTGTTTATCACTGCGACGATTTTGCTACT 60
DB 1 ATGAGAAAGAGAAAGGTTTATCTGTTTATCTGTTTATCACTGCGACGATTTTGCTACT 60
QY 61 GTTGGCTTAAGTATCATCTGCGAGGTGCGAAGAAATTTGATCTGATTTCAAGGA 120
DB 61 GTTGGCTTAAGTATCATCTGCGAGGTGCGAAGAAATTTGATCTGATTTCAAGGA 120
QY 121 ATTGACACAACTGATCTAAGGTTTCTCCAGCAGGGCAGACTGCTGCTGCT 180
DB 121 ATTGACACAACTGATCTAAGGTTTCTCCAGCAGGGCAGACTGCTGCTGCTGCT 180
QY 181 TTTCTGCTGGAATCTGAAATATGAAATCTCCAAAGGTTTGAGAGAGCTTGAAACA 240
DB 181 TTTCTGCTGGAATCTGAAATATGAAATCTCCAAAGGTTTGAGAGAGCTTGAAACA 240
QY 241 GTCCGGGCAAAATTAATCACTCATATTTATCAATTCATGAGCAATTTTGAAGAAACA 300
DB 241 GTCCGGGCAAAATTAATCACTCATATTTATCAATTCATGAGCAATTTTGAAGAAACA 300
QY 301 AAACAGCAGCTGGAAGAAAGAGGGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
DB 301 AAACAGCAGCTGGAAGAAAGAGGGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
QY 361 ATTGTCAGATATGAGGGCGCATGTTAAGTACAGCAACAGCACTGAGAGCTGGAATCC 420
DB 361 ATTGTCAGATATGAGGGCGCATGTTAAGTACAGCAACAGCACTGAGAGCTGGAATCC 420
QY 421 GTGAGACCTTATTTGCCGATATACAGAAATAGATCCAGCTTTTCAAAAAGGGCATCA 480
DB 421 GTGAGACCTTATTTGCCGATATACAGAAATAGATCCAGCTTTTCAAAAAGGGCATCA 480
QY 481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCAATTAGA 540
DB 481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCAATTAGA 540
QY 541 GGCAATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATTAATTAACGGCAAG 600
DB 541 GGCAATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATTAATTAACGGCAAG 600
QY 601 CCTGAGTAAAGTATGATGATGTTGCGCGTGGAAATTTGCAAAAGCGGATGCTGAG 660
DB 601 CCTGAGTAAAGTATGATGATGTTGCGCGTGGAAATTTGCAAAAGCGGATGCTGAG 660
QY 661 AGCAGCTACCGGTTGTATGACAAAGACAGATCTAGCGGTTGCCGATACAGGGCTTGA 720
DB 661 AGCAGCTACCGGTTGTATGACAAAGACAGATCTAGCGGTTGCCGATACAGGGCTTGA 720
QY 721 ACAGGTCCCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTCCCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTGGACGAGCAATATATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGACGAGCAATATATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATTATGAGAAACGCTCCCATTAATAAGGAATGCGGCTCAGGCGCAATCTACTTTC 900
DB 841 TCCGATTATGAGAAACGCTCCCATTAATAAGGAATGCGGCTCAGGCGCAATCTACTTTC 900
```

```
DB 841 TCCGATTATGAGAAACGCTCCCATTAATAAGGAATGCGGCTCAGGCGCAATCTACTTTC 900
QY 901 CAATCTATCATGATATGAGGGGTGGGAGTCTGAGAGACTTACCTTCAATCTGCAACCTTA 960
DB 901 CAATCTATCATGATATGAGGGGTGGGAGTCTGAGAGACTTACCTTCAATCTGCAACCTTA 960
QY 961 TTCACCGAAGCATTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 TTCACCGAAGCATTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 AATGGGCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 AATGGGCTTACACAAAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ACAGTCTTTTCTGCTGCGGGAATGAGAACCGAACCGGCAACCATGATGATGATGATG 1140
DB 1081 ACAGTCTTTTCTGCTGCGGGAATGAGAACCGAACCGGCAACCATGATGATGATGATG 1140
QY 1141 ACAGTAAATATGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 ACAGTAAATATGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 TATGCGCAAAATATCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 TATGCGCAAAATATCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 CGGATCAAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 CGGATCAAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GCACCGGATCTCTCTCTGCGGGAACCATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GCACCGGATCTCTCTCTGCGGGAACCATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TCCATGCTTACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
DB 1381 TCCATGCTTACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
QY 1441 AACGAGGATCAACCAAGGCTTCTTATTAAGGCGGATGATGATGATGATGATGATGAT 1500
DB 1441 AACGAGGATCAACCAAGGCTTCTTATTAAGGCGGATGATGATGATGATGATGATGAT 1500
QY 1501 GACATCGGCTTGGGCTACCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 GACATCGGCTTGGGCTACCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 TCCCTGAAAGTGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 TCCCTGAAAGTGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 TACTGTTTACTGCTATGCTGCGGCAAGCTTTGAAATCTCCCTGCTGATGATGATG 1680
DB 1621 TACTGTTTACTGCTATGCTGCGGCAAGCTTTGAAATCTCCCTGCTGATGATGATG 1680
QY 1681 CCTGAGACCAACCTGCTTCCGTAAGCGTTGCAATGATGATGATGATGATGATGATGAT 1740
DB 1681 CCTGAGACCAACCTGCTTCCGTAAGCGTTGCAATGATGATGATGATGATGATGATGAT 1740
QY 1741 CCAATGCAACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
DB 1741 CCAATGCAACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
QY 1801 GCGCGCAATTAACGTAAGAAATGATTTTATTAATGACACCAAAAGCGGAGCTATCAATT 1860
DB 1801 GCGCGCAATTAACGTAAGAAATGATTTTATTAATGACACCAAAAGCGGAGCTATCAATT 1860
QY 1861 GAGGTACAGGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1861 GAGGTACAGGCTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923
```


Db	1141	ACAGCTAATAATGCAATTAACATTCGAGCTACGGAAAACTTCGCCCAAGCTTTGGATCT	1200
QY	1201	TATGCGACATATCAACCATGTGCAAGTTCTCTTCAAGTGAACGACAAGATGGA	1260
Db	1201	TATGCGACATATCAACCATGTGCAAGTTCTCTTCAAGTGAACGACAAGATGGA	1260
QY	1261	CGAATCAACCGGATGTCAATGACACGGGAAAGTTCAATATGCAAGAAATCTTCTCTT	1320
Db	1261	CGAATCAACCGGATGTCAATGACACGGGAAAGTTCAATATGCAAGAAATCTTCTCTT	1320
QY	1321	GCACCGGATTCCTCTTCTGCGGCAACATGAACATTAATATGCAATATGCAATGCGTGAACG	1380
Db	1321	GCACCGGATTCCTCTTCTGCGGCAACATGAACATTAATATGCAATATGCAATGCGTGAACG	1380
QY	1381	TCGATGGCTACACCGATCGTTGCTGAAAACGTGGCAACGCTTCGTGAGCAATTTTGTGAA	1440
Db	1381	TCGATGGCTACACCGATCGTTGCTGAAAACGTGGCAACGCTTCGTGAGCAATTTTGTGAA	1440
QY	1441	AACAGAGCATACACAAAGGCTTCTCTAATTAAGGGGACATGATGGCGGTGACGCT	1500
Db	1441	AACAGAGCATACACAAAGGCTTCTCTAATTAAGGGGACATGATGGCGGTGACGCT	1500
QY	1501	GACATCGGCTTGGCTACCCGACCGTAAACCAAGATGGGACGATGACATTTGGATAA	1560
Db	1501	GACATCGGCTTGGCTACCCGACCGTAAACCAAGATGGGACGATGACATTTGGATAA	1560
QY	1561	TCCCTGAACGTTGACCTATGTGAAGATCCAGTTCTCTATCCACAGCCAAAAGCGACG	1620
Db	1561	TCCCTGAACGTTGACCTATGTGAAGATCCAGTTCTCTATCCACAGCCAAAAGCGACG	1620
QY	1621	TACTGGTTACTGCTACTACGCGGCAAGCCTTTGAAAATCTCCCTGTATGCTGATGCC	1680
Db	1621	TACTGGTTACTGCTACTACGCGGCAAGCCTTTGAAAATCTCCCTGTATGCTGATGCC	1680
QY	1681	CCTGGACACAAACGCTTCCGTAAACGTTGTCAATGATCTGACCTTTGATTAACGCGT	1740
Db	1681	CCTGGACACAAACGCTTCCGTAAACGTTGTCAATGATCTGACCTTTGATTAACGCGT	1740
QY	1741	CCAAATGGCACACAGTATGTAGAAATGACTTTACTTGCACATCAATGATTAATCGGAT	1800
Db	1741	CCAAATGGCACACAGTATGTAGAAATGACTTTACTTGCACATCAATGATTAATCGGAT	1800
QY	1801	GGCGGCATTAACGTAGAAATGTATTATTAATGACACCAAAACGGGACGTATACAAAT	1860
Db	1801	GGCGGCATTAACGTAGAAATGTATTATTAATGACACCAAAACGGGACGTATACAAAT	1860
QY	1861	GAGGTACAGGCTTATTAAGTACGTACCGGTTGACACAGACCTTCGTGTGGCAATTTGAT	1920
Db	1861	GAGGTACAGGCTTATTAAGTACGTACCGGTTGACACAGACCTTCGTGTGGCAATTTGAT	1920
QY	1921	TAA 1923	
Db	1921	TAA 1923	
RESULT 3			
ADSI14428			
ID ADSI14428 standard; DNA; 1923 BP.			
XX	ADSI14428;		
XX	AC		
XX	30-DEC-2004 (first entry)		
XX	DE	Bacillus alkaline protease KPa3 gene SRO ID NO:2.	
XX	KW	protease; enzyme; alkaline protease; laundry detergent; KPa3; ds; gene.	
XX	OS	Bacillus sp.; KSM-KP43.	
XX	Key	Location/Qualifiers	
XX	FT	1..1923	
XX	CD	/*tag= b	

FT		/product= "alkaline protease"
FT	sig_peptide	1..618
FT		/*tag= a
FT	mat_peptide	619..1920
FT		/*tag= c
XX	PN	
XX	PI	EPI466962-AI.
PD		
PD	13-OCT-2004.	
PF	08-APR-2004; 2004EP-00008604.	
PR	10-APR-2003; 2003JP-00106709.	
XX		
XX	(KAOS) KAO CORP.	
PA		
XX	Okuda M, Sato T, Takimura Y, Sumitomo N, Kobayashi T;	
PI		
DR	WPI: 2004-711313/70.	
DR	P-PsDB; ADS14437, ADS14427.	
XX		
PT	New engineered alkaline protease, useful particularly in laundry	
PT	detergents, comprising specified amino acids at particular positions.	
XX		
PS	Disclosure; SEQ ID NO 2; 31pp; English.	
CC		
CC	The invention relates to a novel alkaline protease. The new alkaline	
CC	protease comprises an amino acid sequence in which one or more amino acid	
CC	residues selected from those located at 7 specific positions within	
CC	ADSI4427, or at positions corresponding to these positions are: position	
CC	15 (histidine), position 16 (threonine or glutamine), position 166	
CC	(glycine), position 167 (valine), position 187 (serine), position 346	
CC	(arginine), and position 405 (aspartic acid). The alkaline protease is	
CC	useful in industry particularly in laundry detergents, but also e.g. in	
CC	fibre modifying agents, leather processing agents, cosmetic compositions,	
CC	bath additives, food-modifying agents, and pharmaceuticals. The present	
CC	sequence encodes the wild-type <i>Bacillus</i> sp. KSM-KP43 alkaline protease,	
CC	KP43.	
SEQ	Sequence 1923 BP; 576 A; 418 C; 475 G; 454 T; 0 U; 0 Other;	
Query Match	100.0%; Score 1923; DB 13; Length 1923;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1923; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 ATGAGAAAGAAGAAAAGGTGTTTATCTGGTTTATCAGTCGACGATTTCGACT	60
Db	1 ATGAGAAAGAAGAAAAGGTGTTTATCTGGTTTATCAGTCGACGATTTCGACT	60
OY	61 GTTGCGTTAAGTAATCCATCTGCAGGTGTCGAAGAAATTTGATCTTGATTTCAAGGA	120
Db	61 GTTGCGTTAAGTAATCCATCTGCAGGTGTCGAAGAAATTTGATCTTGATTTCAAGGA	120
OY	121 ATTGAGCAACAACCTGATGCTAAAGGTTTCTCCAGACGAGGGCAGACTGCTGCTCT	180
Db	121 ATTGAGCAACAACCTGATGCTAAAGGTTTCTCCAGACGAGGGCAGACTGCTGCTCT	180
OY	181 TTTCGTGTGATCTGAAAAATGTGAAAATCCCAAAGAGTTTGACAGAAAGCTTGAACA	240
Db	181 TTTCGTGTGATCTGAAAAATGTGAAAATCCCAAAGAGTTTGACAGAAAGCTTGAACA	240
OY	241 GTCCCGCAATAATTAATCTCATATTAATCCAAATTCAGTGCAATTTTGAAGAACCA	300
Db	241 GTCCCGCAATAATTAATCTCATATTAATCCAAATTCAGTGCAATTTTGAAGAACCA	300
OY	301 AAAACGACGCTGCAAAAAACAGGGGCAGAAAGTTCTGCACTTACATACCAGATTAGCTTAC	360
Db	301 AAAACGACGCTGCAAAAAACAGGGGCAGAAAGTTCTGCACTTACATACCAGATTAGCTTAC	360
OY	361 ATTGTGAGTAGTAGGGCGATGTTTAAGTCAGCAACAAGCACATTGAGCAGCTGGAATCC	420
Db	361 ATTGTGAGTAGTAGGGCGATGTTTAAGTCAGCAACAAGCACATTGAGCAGCTGGAATCC	420

OY	421	GTGGAGCCCTTATTTGCGGATATACGAATGATGCCAGCTTTTCAAAAAGGGGACATCA	480
Db	421	GTGGAGCCCTTATTTGCGGATATACGAATGATGCCAGCTTTTCAAAAAGGGGACATCA	480
OY	481	GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAATAAAGAGGTGCATTTAAGA	540
Db	481	GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAAATAAAGAGGTGCATTTAAGA	540
OY	541	GGCATTCGAACAAAATGGCACAATTCGCAATATGATGTGCTATATATTAACGGCAAG	600
Db	541	GGCATTCGAACAAAATGGCACAATTCGCAATATGATGTGCTATATATTAACGGCAAG	600
OY	601	CCTGAGTATTAAGGTATGATGATGTTGGCGCTGGAATTTGTCAAAGCGGATGTGGCTCAG	660
Db	601	CCTGAGTATTAAGGTATGATGATGTTGGCGCTGGAATTTGTCAAAGCGGATGTGGCTCAG	660
OY	661	AGCAGCTACGGGTTTATGGAACAAGACAGATCTGATCGGTTGCGGATACAGGGCTTGAT	720
Db	661	AGCAGCTACGGGTTTATGGAACAAGACAGATCTGATCGGTTGCGGATACAGGGCTTGAT	720
OY	721	ACAGGTCGCAATGACAGTTGCATGATGAAGCTTTCGGCGGGAATAATTACTGCATTAATAT	780
Db	721	ACAGGTCGCAATGACAGTTGCATGATGAAGCTTTCGGCGGGAATAATTACTGCATTAATAT	780
OY	781	GCAATTGGGACCGGACGAATTAATGCCAATGATACGAATGTGTCAATGTGACATGTGGCTGGC	840
Db	781	GCAATTGGGACCGGACGAATTAATGCCAATGATACGAATGTGTCAATGTGACATGTGGCTGGC	840
OY	841	TCCGATTAAGGAAAAGGCTCCACTAATTAAGGAATGGCGCTCAGGCGAATGTAGCTTTC	900
Db	841	TCCGATTAAGGAAAAGGCTCCACTAATTAAGGAATGGCGCTCAGGCGAATGTAGCTTTC	900
OY	901	CAATCTATCATGATAGCGGTGGGGGACTTGGAGACTACCTTCGATCTGCAACTTA	960
Db	901	CAATCTATCATGATAGCGGTGGGGGACTTGGAGACTACCTTCGATCTGCAACTTA	960
OY	961	TTCAAGCCAAAGCATCAAGTGTGTGTGCGAAGATTCATACAACTCTTGGGAGCAGCAGTG	1020
Db	961	TTCAAGCCAAAGCATCAAGTGTGTGTGCGAAGATTCATACAACTCTTGGGAGCAGCAGTG	1020
OY	1021	AATGAGGCTTACACAACAGATTCAGGAATGTGATGACTATGTGCGGAAAAATGATATNG	1080
Db	1021	AATGAGGCTTACACAACAGATTCAGGAATGTGATGACTATGTGCGGAAAAATGATATNG	1080
OY	1081	ACGATCTTTTGTGCTGCGGGAATGAAGACCGAACGCGGAACCATGATGACACAAGC	1140
Db	1081	ACGATCTTTTGTGCTGCGGGAATGAAGACCGAACGCGGAACCATGATGACACAAGC	1140
OY	1141	ACAGCTTAAATATGCAATTAACAGTCCGAAGCTACGGAATAACCTTCGCGCAAGCTTTGGTCT	1200
Db	1141	ACAGCTTAAATATGCAATTAACAGTCCGAAGCTACGGAATAACCTTCGCGCAAGCTTTGGTCT	1200
OY	1201	TATGCGGACAAATATCAACATGTGCAAGTCTTTCAACGTGACCGGACAAAGATGGA	1260
Db	1201	TATGCGGACAAATATCAACATGTGCAAGTCTTTCAACGTGACCGGACAAAGATGGA	1260
OY	1261	CGGATCAAAACCGGATGTCAATGGCACCGGGAACGTTCAATCTATACGACAGATCTTCTCTT	1320
Db	1261	CGGATCAAAACCGGATGTCAATGGCACCGGGAACGTTCAATCTATACGACAGATCTTCTCTT	1320
OY	1321	GCACCGGATTCCTCTCTTGCGGGAACATGACAGTAAATATGCAATCATGGGTGGAACG	1380
Db	1321	GCACCGGATTCCTCTCTTGCGGGAACATGACAGTAAATATGCAATCATGGGTGGAACG	1380
OY	1381	TCGATGTGTTAACCGGATGTGTGTGGAACGTGGAACGCTTCGGAACATTTTGTGAAA	1440
Db	1381	TCGATGTGTTAACCGGATGTGTGTGGAACGTGGAACGCTTCGGAACATTTTGTGAAA	1440
OY	1441	AAACAGAGCATCAACCAAAAGCTTCTCTAATTAAGCGGACCTAATTTGCGGATGACCT	1500
Db	1441	AAACAGAGCATCAACCAAAAGCTTCTCTAATTAAGCGGACCTAATTTGCGGATGACCT	1500
OY	1501	GACATCGGCTTGTGCTACCGGAACGTTAACGAAGATGGGGAACGATGACATTTGATATA	1560

Db	1501	GACATCGGCCTTGGCTACCCGACGGTAACCAAGATGGGAGCAATGGATPAA	1560
Qy	1561	TCCTGGAAGTTCCTATATGTGACGAATCCAGTCTCTATCCACCAAGCAAAAGCGACG	1620
Db	1561	TCCGTGAACGTTCCTATATGTGAAAGAGTCCAGTCTCTATCCACCAAGCAAAAGCGACG	1620
Qy	1621	TACTCGTTTACTGCTACTCCCGGCAAGCCTTTGAAAATCTCCCTGTATGGTCTGATGCC	1680
Db	1621	TACTCGTTTACTGCTACTGCGGCAAGCCTTTGAAAATCTCCCTGTATGGTCTGATGCC	1680
Qy	1661	CCTCGAGCAACAATGCTTCCTCGTAAAGCTTGTGAATCTGACCTTGTCACTAACCGCT	1740
Db	1661	CCTCGAGCAACAATGCTTCCTCGTAAAGCTTGTGAATCTGACCTTGTCACTAACCGCT	1740
Qy	1741	CCAAATGGCACACAGTATGTAGAAAATGACTTTACTTCGCAATCAATGATTAAGTGGAT	1800
Db	1741	CCAAATGGCACACAGTATGTAGAAAATGACTTTACTTCGCAATCAATGATTAAGTGGAT	1800
Qy	1801	GAGGTACAGGCTTATTAACGTACCGGTGGACCAACAGCCTTTCGTGGCAATTGTGAT	1920
Db	1801	GAGGTACAGGCTTATTAACGTACCGGTGGACCAACAGCCTTTCGTGGCAATTGTGAT	1920
Qy	1861	GAGGTACAGGCTTATTAACGTACCGGTGGACCAACAGCCTTTCGTGGCAATTGTGAT	1920
Db	1861	GAGGTACAGGCTTATTAACGTACCGGTGGACCAACAGCCTTTCGTGGCAATTGTGAT	1920
Qy	1921	TAA 1923	
Db	1921	TAA 1923	
RESULT 4			
ADT49605			
ID	ADT49605	standard; DNA; 1923 BP.	
XX	ADT49605;		
AC	30-DEC-2004	(first entry)	
DT			
XX			
DE	Bacillus alkaline protease encoding DNA.		
XX			
XX	Alkaline protease; detergent; fiber modification; leather processing;		
KW	cosmetic; bath additives; food-modification; pharmaceutical; gene; de.		
XX			
OS	Bacillus sp. KSM-KP43.		
XX			
XX			
Key	Location/Qualifiers		
FT	1..1923		
FT	/*tag= b		
FT	/product= "alkaline protease"		
FT	/note= "SEQ ID 10"		
FT	sig_peptide		
FT	1..618		
FT	/*tag= a		
FT	/note= "signal peptide"		
FT	619..1923		
FT	/*tag= c		
FT	/product= "mature alkaline protease"		
FT	/note= "SEQ ID 1"		
XX			
XX	BP1466970-AL.		
XX			
PD	13-OCT-2004.		
XX			
PF	08-APR-2004; 2004EP-00008605.		
XX			
PR	10-APR-2003; 2003JP-00106708.		
XX			
PA	(KAOS) KAO CORP.		
XX			
PI	Sato T, Okuda M, Koyama S, Izawa Y, Kobayashi T;		
XX	WPI; 2004-711317/70.		

QY 1861 GAGGTACAGGCTTATTAACGTACCGGTGACACAGACCTTCTGTTGGCAATTGTGAAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTACCGGTGACACAGACCTTCTGTTGGCAATTGTGAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923
RESULT 5
AAK37279
ID AAK37279 standard; DNA; 1923 BP.
XX
AC AAK37279;
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.
XX
KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KM washing composition; oxidizing agent; ss.
XX
OS Bacillus sp.
XX
PN MO918218-A1.
XX
PD 15-APR-1999.
XX
PF 07-OCT-1998; J:98MO-JP04528.
XX
PR 07-OCT-1997; 97JP-00274570.
XX
PI (KAOS) KAO CORP.
XX
PI Takaiwa M, Okuda M, Saeiki K, Kubota H, Hltomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
DR WPI, 1999-287736/27.
XX
DR P-PSDB; AAY17091.
XX
PT Alkali protease from Bacillus used in washing powders.
XX
PS Disclosure; Page 63-68; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1923 BP; 578 A; 417 C; 474 G; 454 T; 0 U; 0 Other;
Query Match 99.8%; Score 1919.8; DB 2; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGCGACGATTTGTGACT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGCGACGATTTGTGACT 60
QY GTTGCCTTAAGTATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120
DB 61 GTTGCCTTAAGTATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120

QY 121 ATTACAGCAACAATGATGCTAAAGTTTCTCCAGCGAGGGGCAAGTGGTCTGCTGCT 180
DB 121 ATTACAGCAACAATGATGCTAAAGTTTCTCCAGCGAGGGGCAAGTGGTCTGCTGCT 180
QY 181 TTTCTGGTGAATCTGAAAAATGTGAACCTCCCAAAAGGTTTGGCAAGAAAGCTTGAACA 240
DB 181 TTTCTGGTGAATCTGAAAAATGTGAACCTCCCAAAAGGTTTGGCAAGAAAGCTTGAACA 240
QY 241 GTCCCGGCAAAATATAAATCTCATATATATCAATTCGAATGAGCAATTTTGAAGAAACA 300
DB 241 GTCCCGGCAAAATATAAATCTCATATATATCAATTCGAATGAGCAATTTTGAAGAAACA 300
QY 301 AAACAGCAGCTGGAAGAAACAGGGGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
DB 301 AAACAGCAGCTGGAAGAAACAGGGGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
QY 361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGGACGTTGGAATCC 420
DB 361 ATTGTGAGTATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGGACGTTGGAATCC 420
QY 421 GTGAGCTTATTTGCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGGCAATCA 480
DB 421 GTGAGCTTATTTGCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGGCAATCA 480
QY 481 GAGCTGTAAAGCAGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTTAAGA 540
DB 481 GAGCTGTAAAGCAGTGGCGCTTGATACAAAGCAAGAAATTAAGAGTGCATTTAAGA 540
QY 541 GGCATGAAACAAATGCGACAATTCGCAATTAAGCAATGATGATTAATTAAGGGCAAG 600
DB 541 GGCATGAAACAAATGCGACAATTCGCAATTAAGCAATGATGATTAATTAAGGGCAAG 600
QY 601 CCTGAGTATTAAGTGAATGATGATGTCGCGCTGGAATTTGCAAGCGAGTGGCTCAG 660
DB 601 CCTGAGTATTAAGTGAATGATGATGTCGCGCTGGAATTTGCAAGCGAGTGGCTCAG 660
QY 661 AGCAGCTACGGGTTGATGAGCAAGAGCAGATGTCGAGGTTCCGATACAGGGCTTGAT 720
DB 661 AGCAGCTACGGGTTGATGAGCAAGAGCAGATGTCGAGGTTCCGATACAGGGCTTGAT 720
QY 721 ACAGGTGCAATGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCAATGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTTGGACGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTTGGACGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATTTAGAAACGGCTCCATTAATAAGAAAGGCGCTTCAAGCGAATCTAGCTTTC 900
DB 841 TCCGATTTAGAAACGGCTCCATTAATAAGAAAGGCGCTTCAAGCGAATCTAGCTTTC 900
QY 901 CAATCTATATGATATGCGGTGGGGAATTTGAGAGACTACCTTCAATCTGCAAACTTCA 960
DB 901 CAATCTATATGATATGCGGTGGGGAATTTGAGAGACTACCTTCAATCTGCAAACTTCA 960
QY 961 TTCAGCAAGCAATACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1020
DB 961 TTCAGCAAGCAATACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1020
QY 1021 AATGGGGCTTACACACAGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGGCTTACACACAGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCTTTTGGCTGGCGGGAATGAAGGACGGAACGGGGAACCACTGATGACCAAGGC 1140
DB 1081 ACGATCTTTTGGCTGGCGGGAATGAAGGACGGAACGGGGAACCACTGATGACCAAGGC 1140
QY 1141 ACAGCTAAATAATGCAATTAACATGCGAGCTACGGAACCTCCGCCAAGCTTTGGGTCT 1200
DB 1141 ACAGCTAAATAATGCAATTAACATGCGAGCTACGGAACCTCCGCCAAGCTTTGGGTCT 1200

```

QY 1201 TATCGGAGCAATATCAACGATGAGCACTTCTCTCACTGAGACCAAGATGGA 1260
DB 1201 TATCGGAGCAATATCAACGATGAGCACTTCTCTCACTGAGACCAAGATGGA 1260
QY 1261 CGGATCAAAACCGGATGATGAGCAACGGGAAACGTTCACTATCAAGATCTTCTCT 1320
DB 1261 CGGATCAAAACCGGATGATGAGCAACGGGAAACGTTCACTATCAAGATCTTCTCT 1320
QY 1321 GCACCGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
DB 1321 GCACCGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
QY 1381 TCCATGAGCTACACGATGCTGCTGAGAAAGTGGACAGCTTGATGAGATTTTGAAA 1440
DB 1381 TCCATGAGCTACACGATGCTGCTGAGAAAGTGGACAGCTTGATGAGATTTTGAAA 1440
QY 1441 AACAGAGGATCAACACCAAGCCTTCTATTTAAAGCGGACATGATGCGGTGAGCT 1500
DB 1441 AACAGAGGATCAACACCAAGCCTTCTATTTAAAGCGGACATGATGCGGTGAGCT 1500
QY 1501 GACATGCGCTTGGCTTACCCGAAACGGTAACCAAGATGAGGACGAGTGAATTGATTA 1560
DB 1501 GACATGCGCTTGGCTTACCCGAAACGGTAACCAAGATGAGGACGAGTGAATTGA 1560
QY 1561 TCCCTGAACGTTGCTATGTAAGAGTCAAGTCTCTATTCACGACCAAAAGGAGAG 1620
DB 1561 TCCCTGAACGTTGCTATGTAAGAGTCAAGTCTCTATTCACGACCAAAAGGAGAG 1620
QY 1621 TACTGCTTTACTGCTACTGCTGCGGCAAGCCTTTGAAAATCTCCCTGATGCTGATGCC 1680
DB 1621 TACTGCTTTACTGCTACTGCTGCGGCAAGCCTTTGAAAATCTCCCTGATGCTGATGCC 1680
QY 1681 CCGTGGAGGACAACTGCTTCCGTAAGCTTGTCAAGATGCTGACCTTGTCAATCCGT 1740
DB 1681 CCGTGGAGGACAACTGCTTCCGTAAGCTTGTCAAGATGCTGACCTTGTCAATCCGT 1740
QY 1741 CCAAAATGAGCAACAGTATGAGAAATGACTTACTTCCCATCAATGATTAATCGGAT 1800
DB 1741 CCAAAATGAGCAACAGTATGAGAAATGACTTACTTCCCATCAATGATTAATCGGAT 1800
QY 1801 GCGCCGCAATACGTAAGAAATGATTTATTAATGACACCAAAAGCGGAGCTATCAAT 1860
DB 1801 GCGCCGCAATACGTAAGAAATGATTTATTAATGACACCAAAAGCGGAGCTATCAAT 1860
QY 1861 GAGTACAGGCTTATTAAGTACCGGTTGACCAAGACCTTCTCGTTGCAATTTGAA 1920
DB 1861 GAGTACAGGCTTATTAAGTACCGGTTGACCAAGACCTTCTCGTTGCAATTTGAA 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

```

RESULT 6
AAK37277
ID AAK37277 standard; DNA; 1920 BP.

```

XX AAK37277;
AC 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease encoding DNA.
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidizing agent; ss.
XX
OS Bacillus sp.
XX
PN MO9918218-A1.
XX
PD 15-APR-1999.
XX

```

```

PF 07-OCT-1998; 98MO-JP004528.
XX
XX PR 07-OCT-1997; 97JP-00274570.
XX
XX (KAOS ) KAO CORP.
PI Takaiwa M, Okuda M, Sasaki K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M;
DR WPI; 1999-287736/27.
DR P-PSDB; AAY17087, AAY17089.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Disclosure; Page 53-58; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease encoding
CC DNA. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 1920 BP, 576 A, 411 C, 473 G, 460 T, 0 U, 0 Other;
Query Match 76.8%; Score 1477.2; DB 2; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
QY 10 AAGAAAAAGGCTTTTATCTGTTTATCACTGACGAGATTTGTGACTGTTCGTTA 69
DB 7 AAGAAAGAGGCTTTTATCTGTTTATCACTGACGAGATTTGTGACTGTTCGTTA 66
QY 70 AGTATCCATCTGACGAGTGTGCAAGGAATTTGATCTGATTTCAAGGAATTCAAGCA 129
DB 67 AACATTCCTCGCTGTGTGATGACAGACTTTTATCTGATTTTAAAGGAATTCANCA 126
QY 130 ACAACTAATGCTAAAGTTTCTCCAGCAGGAGCACTGTGTGCTTTTCTGTG 189
DB 127 ACAACCAATGCTAATGCTTTCTCCAAACAGGCAAAACAGGTGGCTGCTTTCTGTG 186
QY 190 GAATCGAAAAATGGAATCCCAAAAGCTTTGCAAGAAAGCTTGAACAGTCCCGCA 249
DB 187 GAGTCTAAAAATGGAATCTTTAAAGATGCTTAAAGAACTTGAACAGTACCGCA 246
QY 250 AATTAATAAATCCATATTATTCATATGACCAATTTTAAAGAAACAAACACAG 309
DB 247 AATTAATAAATCCATATTGTCATATGAGCCCAATTTTAAAGAAACAAACAGAG 306
QY 310 CTGAAAAAACAAGGAGCAAGATTTCTGACTACATACCTGATTAATGCTTAC 369
DB 307 CTAGAGCAAACTGAGAGCAAGATTTCTGACTACATACCTGATTAATGCTTAC 366
QY 370 TATGAGGAGATGTTAATGACGAACAAGCACTTGAAGCACTGAGAACTCCGTGAGCT 429
DB 367 TATGAGGAGATGTTAATGACGAACAAGCACTTGAAGCACTGAGAACTCCGTGAGCT 426
QY 430 TATTTGCGATATACAGAAATGATCCCACTTTTCAAAAGGAGCATGAGACTTGT 489
DB 427 TACTTGCGAAATATCAAAATAGATCCCACTTTTCAAAAGGAGCATGAGACTTGT 486
QY 490 AAGCAGTGGCGCTTGTATCAAAAGCAAAATTAAGAGGTGCAATTAAGAGCATGAA 549
DB 487 AAGCGTTGGCGCTTGTATCAAAAGCAAAATTAAGAGGTGCAATTAAGAGCATGAG 546

```

QY 550 CAATCGCACAAATTCGCAATTAAGCATATGCTATATATTAAGCGAAAGCCTGATAT 609
 DB 547 GAATGCTCAGTACGTACGACGAATGAGTCCATTTATTAAGCGAAAGCCTGATAT 606
 QY 610 AAGGTGATGATATGTTGCGCGTGAATTTGTCAAAGCGGATGTGCTCAGAGCAGTAC 669
 DB 607 AAGGTGATGATATGTTGCGCGTGAATTTGTCAAAGCGGATGTGCTCAGAGCAGTAC 666
 QY 670 GGGTGTGATGACAAAGACAGATCTGTAGCGGTTGCCCATACAGGGCTTGATACAGTCCG 729
 DB 667 GGGTGTGATGACAAAGACAGATCTGTAGCGGTTGCCCATACAGGGCTTGATACAGGAGA 726
 QY 730 AATGACAGTTCGATGACATGAAGCTTCGCGGGGAAATTAAGTCTGATATATGATGCGA 789
 DB 727 AAGCAGAGTTCGATGACATGAAGCTTCGCGGGGAAATTAAGTCTGATATATGACACTGG 786
 QY 790 CGGACGATATATGCGCAATGATACGATGCTCATGATGATGCGATGTGCTGCGATTA 849
 DB 787 CGGACGATATATGCGCAATGATACGATGCTCATGATGATGCGATGTGCGATTA 846
 QY 850 GGAACCGGCTCCATTAATAAGGAATGCGGCTCAGGCGCAATCTAGTCTTCCATCTATC 909
 DB 847 GGAATATGCGCGCAAGATTAAGGAATGCGACCTCAAGCGAATCTGTTTTCATATCATC 906
 QY 910 AATGATAGCGGTGGGGAATTTGAGAGACTACCTTCCGAACTTGCAAACTTATTCAAGCAA 969
 DB 907 AATGATAGCGGTGGGGAATTTGAGAGACTACCTTCCGAACTTATTCAAGCAA 966
 QY 970 GCATACAGTCTGTGTCGCAAAATTCATACAACTCTGCGGAGACAGCATGAAATGGGCT 1029
 DB 967 GCATACAGTCTGTGTCGCAAAATTCATACAACTCTGCGGAGACAGCATGAAATGGGCT 1026
 QY 1030 TACACACAGATTCGCAAAATTTGATGATGATGCTATGTCGCGCAAAATGATATGACATCTT 1089
 DB 1027 TACACACAGATTCGCAAAATTTGATGATGATGCTATGTCGCGCAAAATGATATGACATCTT 1086
 QY 1090 TTGCTGCGCGGAAATGAGACCGAAACGCGGAAACCATGATGACACGAGCAAGCTTAA 1149
 DB 1087 TTGCTGCGCGGAAATGAGACCGAAACGCGGAAACCATGATGACACGAGCAAGCTTAA 1146
 QY 1150 AATGCAATTAAGTCTGAGCTACGCAAAACCTCGCGCAAGCTTGTGCTTATGCGGAC 1209
 DB 1147 AATGCAATTAAGTCTGAGCTACGCAAAACCTCGCGCAAGCTTGTGCTTATGCGGAC 1206
 QY 1210 AATATCAACATGTTGCGACAGTCTTCTTCACTGTCGACCGCAAAAGATGAGCGATCAA 1269
 DB 1207 AATATCAACATGTTGCGACAGTCTTCTTCCGTCGCGCGCAAAAGATGAGCGATCAA 1266
 QY 1270 CCGGATGTATGTCACCGGGAAGTCTATCTATGACGAGATCTTCTTGTGACCGGAT 1329
 DB 1267 CCGGATGTATGTCACCGGGAAGTCTATCTATGACGAGATCTTCTTGTGACCGGAT 1326
 QY 1330 TCTCTCTTGTGCGCAACATGACAGTAAATATGATATACATGAGTGTGAAAGCTCATGCT 1389
 DB 1327 TCTCTCTTGTGCGCAACATGACAGTAAATATGATATGAGTGTGAAAGCTCATGCT 1386
 QY 1390 ACACCGATGTTGTCGAAAGCTGTGCAACGCTTGTGAGCATTTTGTGAAAGCAAGGC 1449
 DB 1387 ACACCGATGTTGTCGAAAGCTGTGCAACGCTTGTGAGCATTTTGTGAAAGCAAGGC 1446
 QY 1450 ATACACGAAAGCTTCTGATTAATAAGGCGCATGATGCGGTCAGCTGACATTCGCG 1509
 DB 1447 ATACACGAAAGCTTCTGATTAATAAGGCGCATGATGCGGTCAGCTGACATTCGCG 1506
 QY 1510 CTGTGCTACCGCAACGATACCAAGATGCGGAGCATGATGATTAATCCCTGAAC 1569
 DB 1507 CTGTGCTACCGCAACGATACCAAGATGCGGAGCATGATGATTAATCCCTGAAC 1566
 QY 1570 GTTGTCTATGTAAGAGTCACTTCTATTCACCAAGCAAAAGCGATCTGCTT 1629
 DB 1567 GTTGTCTATGTAAGAGTCACTTCTATTCACCAAGCAAAAGCGATCTGCTT 1626
 QY 1630 ACTGTACTGCGCGCAAGCTTTGAAATCTCCTGTGATGCTATGCCCTGTGAGC 1689

DB 1627 ACTGTACTGCGCGCAAGCTTTGAAATCTCCTGTGATGCTATGCTGAGTCCCTGTGAGC 1686
 QY 1690 ACACTGCTTCCGTAAACGCTTGTCAATGATGTGACCTTGTCAATACGCTCCAAATGCG 1749
 DB 1687 ACTGTACTGCTTCCGTAAACGCTTGTCAATGATGTGACCTTGTCAATACGAGCAACGGA 1746
 QY 1750 ACACAGTACTAGGAATGATCTTACTTCCGCTTACATGATATCTGGATGTGCGCGAAT 1809
 DB 1747 ACAGATATGTCGGAATGATCTTCTGACGACCAATTTGACATTAATCTGGATGTGCGCAAT 1806
 QY 1810 AACGTAGAAATATGATTAATTAATGACCAACAAAGCGGAGTATACATTTGAGGTACAG 1869
 DB 1807 AACGTAGAAATATGATTAATTAATGACCAACAAAGTGAACATATACATTTGAGGTACAA 1866
 QY 1870 GCTTATTAACGTACCGGTGAGCAACAGACCTTCTGTTGCAATTTGAAATTA 1923
 DB 1867 GCATATTAATGTCGCGGTGAGCAACAACTTCTGTTGCAATTTGAAATTA 1920

RESULT 7
 ADY33779
 ID ADY33779 standard; DNA; 1305 BP.
 XX
 AC ADY33779;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE Bacillus species alkaline protease coding sequence.
 XX
 KW mutagenesis; alkaline protease; detergent; laundry; cosmetic; food;
 KW bleaching detergent; gene; ds.
 XX
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1305
 FT /*tag= a
 FT /product= "alkaline protease"
 FT /note= "no start codon given"
 FT /partial
 XX
 PN EP1347044-A2.
 XX
 PD 24-SEP-2003.
 XX
 PF 21-MAR-2003; 2003BP-00006472.
 XX
 PR 22-MAR-2002; 2002JP-00081428.
 PR 06-JUN-2002; 2002JP-00165987.
 PR 18-OCT-2002; 2002JP-00304230.
 PR 18-OCT-2002; 2002JP-00304231.
 XX
 PA (KAOS) KAO CORP.
 PA
 XX
 PI Okuda M, Sato T, Saito K, Sumitomo N, Izawa Y, Saeki K;
 PI Kobayashi T, Nomura M;
 DR
 DR MPI; 2003-846540/79.
 DR P-PSDB; ADY33778.
 PT
 PT New alkaline protease having specific amino acid residue at a specific
 PT position of its amino acid sequence, useful for producing detergent
 PT compositions, laundry detergent, fiber modifiers, leather-treating agents
 PT or pipe cleaners.
 PS
 PS Disclosure; SEQ ID NO 2; 31pp; English.
 CC
 CC The invention relates to an alkaline protease having a fully defined
 CC sequence of 434 amino acids (I) given in the specification, or an amino
 CC acid sequence at least 80% homology with (I), where an amino acid residue
 CC at position 65, 101, 163, 170, 171, 273, 320, 359 or 387 of (I) is
 CC selected from 16 amino acid residues. The amino acid residues at the

corresponding positions are selected from: position 65: proline; position 101: asparagine; position 163: histidine, aspartic acid, phenylalanine, lysine, asparagine, serine, isoleucine, leucine, glutamine, threonine or valine; position 170: valine or leucine; position 171: alanine, glutamic acid, glycine or threonine; position 273: isoleucine, glycine or threonine; position 320: phenylalanine, valine, threonine, leucine, isoleucine or glycine; position 359: serine, leucine, valine, isoleucine or glutamic acid; and position 387: alanine, lysine, glutamine, glutamic acid, arginine or histidine. The alkaline protease is useful for the production of a detergent composition, such as laundry detergent, fiber conditioners, leather-treating agents, cosmetic compositions, bath additives, food modifiers and pharmaceutical compositions. The alkaline protease may also be used as bleaching detergent, hard surface cleansing detergent, pipe cleaner, artificial tooth cleaner, and as a sterilizing cleanser for medical tools. The new alkaline protease has a more potent proteolytic capacity, exhibiting excellent detergency for the removal of a complex stain, and has high secretion capacity. This sequence corresponds to the *Bacillus* sp. alkaline protease coding sequence.

Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Query Match 67.9%; Score 1305; DB 11; Length 1305;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACGGGTTGAT 678
DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACGGGTTGAT 60
QY 679 GGAAGAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTGCGATGACGT 738
DB 61 GGAAGAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTGCGATGACGT 120
QY 739 TCGATGCAAGAGCCTTCCGCGGAAATTTATGCTATATATGCAATTTGGAGAGAGAT 798
DB 121 TCGATGCAAGAGCCTTCCGCGGAAATTTATGCTATATATGCAATTTGGAGAGAGAT 180
QY 799 AATGCCAATGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 181 AATGCCAATGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 859 TCCACTAATTAAGAAATGAGCGCTCAGCGGAAATTAAGTCTTCATATGATGATGATGAT 918
DB 241 TCCACTAATTAAGAAATGAGCGCTCAGCGGAAATTAAGTCTTCATATGATGATGATGAT 300
QY 919 GGTGGGGGACTGAGGAGTACCTTTCGATCTGCAAACTTATTCAGCCATACATACAT 978
DB 301 GGTGGGGGACTGAGGAGTACCTTTCGATCTGCAAACTTATTCAGCCATACATACAT 360
QY 979 GCTGTGTCAGAAATTCATCAAACTCTTGGGAGAGCAGTGAATGGGCTTACACAACA 1038
DB 361 GCTGTGTCAGAAATTCATCAAACTCTTGGGAGAGCAGTGAATGGGCTTACACAACA 420
QY 1039 GATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
DB 421 GATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1099 GGAATGAAAGACGAAAGCGGGAACATCACTGAGACCAAGAGAGAGAGAGAGAGAGAG 1158
DB 481 GGAATGAAAGACGAAAGCGGGAACATCACTGAGACCAAGAGAGAGAGAGAGAGAGAG 540
QY 1159 ACAAGTCGAGTACCGAAAACTCCGCCCAAGCTTTGGGTCTTATGCGGACAAATATCAC 1218
DB 541 ACAAGTCGAGTACCGAAAACTCCGCCCAAGCTTTGGGTCTTATGCGGACAAATATCAC 600
QY 1219 CATGTGCAAGATTCTTTCAGTGGAGCCGACAAAGATGAGCGGATCAAAACCGAGTGC 1278
DB 601 CATGTGCAAGATTCTTTCAGTGGAGCCGACAAAGATGAGCGGATCAAAACCGAGTGC 660
QY 1279 ATGGACCGGGAACGTTCACTATCAGAGAGATCTTCTTGGACCGGATTCCTCCCTTC 1338
DB 661 ATGGACCGGGAACGTTCACTATCAGAGAGATCTTCTTGGACCGGATTCCTCCCTTC 720

```

```

QY 1339 TGGGGAAACATGACAGTAAATATGATATGAGGTGAAAGCTCCATGCTTACCCGATC 1398
DB 721 TGGGGAAACATGACAGTAAATATGATATGAGGTGAAAGCTCCATGCTTACCCGATC 780
QY 1399 GTTGTGGAACGTTGACAGAGTCTGAGTATGATGATGATGATGATGATGATGATGATGAT 1458
DB 781 GTTGTGGAACGTTGACAGAGTCTGAGTATGATGATGATGATGATGATGATGATGATGAT 840
QY 1459 AAGCTTCTCTATTAAGAGCGGAGTATGCGGAGTACGATGAGCTGAGCTTGGCTAC 1518
DB 841 AAGCTTCTCTATTAAGAGCGGAGTATGCGGAGTACGATGAGCTGAGCTTGGCTAC 900
QY 1519 CCGAAGCGTAACCAAGATGAGGAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1578
DB 901 CCGAAGCGTAACCAAGATGAGGAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1579 GTGAACGAGTCAAGTCTCTATCCAGCAGCCAAAAGAGAGAGTACGTTTACTGCTACT 1638
DB 961 GTGAACGAGTCAAGTCTCTATCCAGCAGCCAAAAGAGAGAGTACGTTTACTGCTACT 1020
QY 1639 GCGGCAAGCCTTGAAGAAATCTCCCTGATATGATGATGATGATGATGATGATGATGATGAT 1698
DB 1021 GCGGCAAGCCTTGAAGAAATCTCCCTGATATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1699 TCCGTAAGCTTGTCAATGATCTGAGCCTTGTCAATGATGATGATGATGATGATGATGATGAT 1758
DB 1081 TCCGTAAGCTTGTCAATGATCTGAGCCTTGTCAATGATGATGATGATGATGATGATGATGAT 1140
QY 1759 GTAGAAATGACCTTACTTCCCATACATGATGATGATGATGATGATGATGATGATGATGAT 1818
DB 1141 GTAGAAATGACCTTACTTCCCATACATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1819 AATGATTTATTAATGACCAACCAAGCGGAGCGTATCAATTAAGTACAGCTTATAC 1878
DB 1201 AATGATTTATTAATGACCAACCAAGCGGAGCGTATCAATTAAGTACAGCTTATAC 1260
QY 1879 GTACCGGTTGAGCAGACGACCTTCTCGTTGCAATTTGAAATTA 1923
DB 1261 GTACCGGTTGAGCAGACGACCTTCTCGTTGCAATTTGAAATTA 1305

```

```

RESULT 8
ADL25803
ID ADL25803 standard; DNA; 1305 BP.
XX
AC ADL25803;
XX
DT 20-MAY-2004 (first entry)
XX
DE Bacillus alkaline protease coding sequence.
XX
KM alkaline protease; washing agent; gene; ds; enzyme.
XX
OS Bacillus sp.
XX
FH Key
XX CDS 1..1305
XX FT location/Qualifiers
XX FT /*tag=
XX FT /partial
XX FT /product= "Bacillus alkaline protease"
XX FT /note= "No start codon is given"
XX
XX JP2004008085-A.
XX
XX 15-JAN-2004.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX 06-JUN-2002; 2002JP-00165950.
XX
XX (KAOS ) KAO CORP.
XX
XX WPI; 2004-094297/10.
DR

```


DR P-PSDB; ADL25802.

XX Novel mutant alkaline protease produced by substituting the amino acid
PT residues useful as washing agent.

XX Disclosure; SEQ ID NO 2; 21bp; Japanese.

XX The invention comprises a mutant *Bacillus* alkaline protease which
CC contains substitutions at positions 163, 170 and 434. The mutant alkaline
CC protease is useful as a washing agent. The present DNA sequence encodes a
CC *Bacillus* alkaline protease of the invention.

XX Sequence 1305 BP; 366 A; 307 C; 331 G; 301 T; 0 U; 0 Other;

Query Match 67.9%; Score 1305; DB 12; Length 1305;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGCTACGGGTTGTAT 678
DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGCTACGGGTTGTAT 60
QY 679 GGAACAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGTTCGCAATGACAGT 738
DB 61 GGAACAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGTTCGCAATGACAGT 120
QY 739 TCGATGCAATGAAGCTTTCCGCGGAAAATTACTGCAATTATGCAATTGGGACGAGCAAT 798
DB 121 TCGATGCAATGAAGCTTTCCGCGGAAAATTACTGCAATTATGCAATTGGGACGAGCAAT 180
QY 799 AATGCCAATGATACGATGATGTCATGGTAGCGCATGTGGCTGGCTCCGATTTAGGAAACGCG 858
DB 181 AATGCCAATGATACGATGATGTCATGGTAGCGCATGTGGCTGGCTCCGATTTAGGAAACGCG 240
QY 859 TCCACTAATTAAGAAATGGCGCTCAGCGCAATCTAATCTTCCATCTATCATGATAGC 918
DB 241 TCCACTAATTAAGAAATGGCGCTCAGCGCAATCTAATCTTCCATCTATCATGATAGC 300
QY 919 GGTGGGGGACTTGGAGAGTAACTTGTGATCTGCAAACTTATTACGCCAAGATACAGT 978
DB 301 GGTGGGGGACTTGGAGAGTAACTTGTGATCTGCAAACTTATTACGCCAAGATACAGT 360
QY 979 GCTGTGCGCAAAATTCATACAACTCTGGGAGAGCAGTGAATGGGCTTACACAAAC 1038
DB 361 GCTGTGCGCAAAATTCATACAACTCTGGGAGAGCAGTGAATGGGCTTACACAAAC 420
QY 1039 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATCTCTTTTGGCTGCC 1098
DB 421 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATCTCTTTTGGCTGCC 480
QY 1099 GGGAAATGAAGACCGAAGCGCGGAACATCATGTCACACGCGCAGACTTAAATGCAAT 1158
DB 481 GGGAAATGAAGACCGAAGCGCGGAACATCATGTCACACGCGCAGACTTAAATGCAAT 540
QY 1159 ACAAGTCGAGGCTACCGGAAAACCTCCGCGCAAGCTTTGGGCTTATAGCGGCAATATCAAC 1218
DB 541 ACAAGTCGAGGCTACCGGAAAACCTCCGCGCAAGCTTTGGGCTTATAGCGGCAATATCAAC 600
QY 1219 CATGTGCGCAGATTCTTCACTGAGCCGACAAAGATGAGCGATCAACCGGATGTC 1278
DB 601 CATGTGCGCAGATTCTTCACTGAGCCGACAAAGATGAGCGATCAACCGGATGTC 660
QY 1279 AATGGCACCGGGAAAGTTCTATCTATGACAAAGATCTTCTTGGACCGGATTCCTCTTC 1338
DB 661 AATGGCACCGGGAAAGTTCTATCTATGACAAAGATCTTCTTGGACCGGATTCCTCTTC 720
QY 1339 TGGGCGAATCATGACGATTAATATGATCATGGGGAACGTCATGCTACACCGATC 1398
DB 721 TGGGCGAATCATGACGATTAATATGATCATGGGGAACGTCATGCTACACCGATC 780
QY 1399 GTTGTGGAACGATGACAGCTTCTGTGAGCAATTTGTGAAAAACAGAGGATCACACCA 1458
DB 781 GTTGTGGAACGATGACAGCTTCTGTGAGCAATTTGTGAAAAACAGAGGATCACACCA 840

```

```

QY 1459 AAGCTTCTCTATTAAAGGGGACATGATTCGGGTCAGCTGACATCGGCTTGGCTAC 1518
DB 841 AAGCTTCTCTATTAAAGGGGACATGATTCGGGTCAGCTGACATCGGCTTGGCTAC 900
QY 1519 CCGAAGCGTAAACCAAGATGGGAGAGTGAATTTGATTAATCCCTGAACGTTGCTAT 1578
DB 901 CCGAAGCGTAAACCAAGATGGGAGAGTGAATTTGATTAATCCCTGAACGTTGCTAT 960
QY 1579 GTGAACGATGCAAGTTCTCTATCAGCAGCAGCAAAAGCGACGTAATCTGTTACTGCTACT 1638
DB 961 GTGAACGATGCAAGTTCTCTATCAGCAGCAGCAAAAGCGACGTAATCTGTTACTGCTACT 1020
QY 1639 GCCGGCAAGCTTTGAAATTCCTCCGTGATGTGTCATGTCGCTCGAGACACATGCT 1698
DB 1021 GCCGGCAAGCTTTGAAATTCCTCCGTGATGTGTCATGTCGCTCGAGACACATGCT 1080
QY 1699 TCCGTAAAGCTTGTCAATGATCTGACCTTGTCAATTAACGCTCAATGAGCACAGTAT 1758
DB 1081 TCCGTAAAGCTTGTCAATGATCTGACCTTGTCAATTAACGCTCAATGAGCACAGTAT 1140
QY 1759 GTAGAAATGACTTTAATTCCGCAATCAATGATTAATGAGATGCGGCAATTAAGTAA 1818
DB 1141 GTAGAAATGACTTTAATTCCGCAATCAATGATTAATGAGATGCGGCAATTAAGTAA 1200
QY 1819 AATGATTTTATTAATGACCAACAAGCGGAGCTATTAATGAGGTAACGCTTATAC 1878
DB 1201 AATGATTTTATTAATGACCAACAAGCGGAGCTATTAATGAGGTAACGCTTATAC 1260
QY 1879 GTACGGGTTGAGACACAGACCTTCTGTTGGCAATTGTGAATTAA 1923
DB 1261 GTACGGGTTGAGACACAGACCTTCTGTTGGCAATTGTGAATTAA 1305

```

```

RESULT 9
AAV82382
ID AAV82382 standard; DNA; 3003 BP.
XX
AC AAV82382;
XX
DT 12-APR-1999 (first entry)
XX
DE Bacillus Jp170 protease gene.
XX
KW Protease; detergent; surfactant; leather processing; deblotting;
XX flavour; ss.
XX
OS Bacillus sp.
XX
FH Key Location/Qualifiers
FT CDS 846..2771
FT sig_peptide 846..944
FT mat_peptide 1470..2768
FT tag = c
XX
XX WO9856927-A2.
XX
XX 17-DEC-1998.
XX
XX 09-JUN-1998; 98WO-US012005.
XX
XX 12-JUN-1997; 97US-00873479.
XX
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX
XX Sloma A, Christenson L;
XX
XX MPI; 1999-080908/07.
XX
XX P-PSDB; AAV89547.
XX
XX Novel protease from Bacillus subtilis LC20 - useful in laundry and

```

dishwashing detergents and for leather processing.

Claim 11; Page 52-53; 77pp; English.

This nucleotide sequence encodes a novel protease (see AAV82382) of *Bacillus* sp. JP170 (NCIB 12513). The sequence in plasmid p170BAN is contained in *Bacillus subtilis* LC20 NRRL B-21680. The protease gene was isolated from chromosomal DNA of JP170 following preparation of probes based on protease N-terminal and internal peptides (see AAV8549-50), screening of chromosomal libraries, isolation of the 3' end of the gene by inverse PCR (see AAV82410-11), reconstruction of 5' and 3' ends and PCR amplification (see AAV82412-16). Claimed recombinant host cells can be used in a method for producing the protease. The protease is used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for debittering and enhancing the degree of hydrolysis of protein hydrolysates, for flavour development through hydrolysis of proteins, degradation of undesired peptides and in enzymatic synthesis of peptides. It has enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the peroxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins

Sequence 3003 BP; 982 A; 504 C; 645 G; 872 T; 0 U; 0 Other;

Query Match 100.0%; Score 1166.4; DB 2; Length 3003;

Best Local Similarity 76.0%; Pred. No. 0;

Matches 1453; Conservative 0; Mismatches 456; Indels 3; Gaps 1;

12 GAAAAAGTGTCTTTATCTGTTTATCAGCTGACGAGATTGTCGATGTTGCGTTAAG 71
Db 863 GAAGAGGGTGTCTTTATCTGTTTATCAGCTGACGATGTTGTCGATGTTGCGTTAAG 922
Qy 72 TAATCATCTGACAGGAGTGTGCAAGAAATTTGATCTGATTTCAAGAAATTCAGACAC 131
Db 923 CAGTCTCTTCTATCTATTTGGGCGCAACATTTGAATTTGACCTTTAAGGGGATGAACACT 982
Qy 132 AACTGATGTCTAAGGTTTCTCCAGCAGGGGCGACGCTGCTGCTGCTTTCTGATGGA 191
Db 963 TACGCTAAGAGAGGCTGCCACACAGAGAGAAACGGAAAGGACCTTTCTTGTTAA 1042
Qy 192 ATCTGAAAAATGGAACCTCCCAAAAGGTTGCAAGAAAGCTTGAACAGTCCGGCAAA 251
Db 1043 CTCTAAAAATGGAATTCCTCAAGAGTATTCAAAAGAACTAGAAAGTGTCCAGCGGA 1102
Qy 252 TAATTAACCTCATATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 311
Db 1103 TTAACAGCTATATATCTGTTCAATTTGACGACCTATTTTGAAGAAACGCAACTTCAACT 1162
Qy 312 GGAAGAAAACAGGGGCAAAAGATTCGACTACATACCTGATTAATGCTTACATTTGCGATGA 371
Db 1163 AGAGAAAGCGGAGCGAAATTTCTGATTAACATACAGATTTACGTTTATTTGCGATTA 1222
Qy 372 TGAAGGCGCATGTTAAGTCAAGCAAGCAACCTTGAAGCACTGCGAATCCGTAAGCCTTGA 431
Db 1223 TGAATGGGATGTTAAGGCGCGTAACCTGACATTTGCGATTTGGAATCGTTGAACCTATA 1282
Qy 432 TTGGCGATATACAGATAGATCCCGACGTTTTCACAAAAGGGGCGATCAGAGCTTGTAA 491
Db 1283 TTTACCTTTAT 1342
Qy 492 AGCAGTGGCGCTTGAATCAAAAGCAAGAAATTAAGAGGTGCAATTAAGAGCATCGAAC 551
Db 1343 AACAGTGGCTTGAATTAAGAAAGCAAGAAATTAAGAGGTGCAATTAAGAGGTGGAACA 1402
Qy 552 AATGCAACAATTCGCAATTAAGCAATGATGCTATATATTAACGGCAAGGCTGAGATTA 611
Db 1403 AATGCCCCAATAATGCAAAATTAAGATGATTTAATACCTAACCCCAAGCCTGAAATGCA 1462
Qy 612 GGTGATGATGATGTTGGCGCTGGAATTTGCAAAAGCGGATGCTGAGACAGCTACCG 671
Db 1463 AGTTTGAATGACGTGGCGCTGCGATTTGTAAAGAGACGTCGCAAAATTAATCTTTGG 1522

Qy 672 GTTGTATGACAAAGACAGATGCTGACGGTTCGATACAGGGCTTGTATACAGTTCGCA 731
Db 1523 CTTATATATGACAAAGACAGATGCTGACGGTTCGATACAGGGCTTGTATACAGGAA 1582
Qy 732 TGAAGTTCGATGATGAAAGCCTTCGCGGGGAAATTTCTGATATATATATATATATATAT 791
Db 1583 TGACAGTTCGATGATGAAAGCCTTCGCGGGGAAATTTCTGATATATATATATATATATAT 1642
Qy 792 GACGAATATATGCAATGATGCAATGATGATGATGATGATGATGATGATGATGATGATG 851
Db 1643 AACGAATATATGCAATGATGCAATGATGATGATGATGATGATGATGATGATGATGATG 1702
Qy 852 AAACGGCTCCATTAATAAGAAATGAGCGCTTCAGCGCAATTAAGTCTTCAATCTATCAT 911
Db 1703 AAATG---CTACAAATTAAGAAATGAGCGCTTCAGCGCAATTAAGTCTTCAATCTATCAT 1759
Qy 912 GGATAGCGGTGGGGGACTTGGAGAGCTACTTCGAAATCTGCAAACTTAATTCAGCGAGC 971
Db 1760 GGATAGGTGTGAGAGGCTGGAGAGCTACTTCGAAATCTGCAAACTTAATTCAGCGAGC 1819
Qy 972 ATACAGTCTGTGTCAGAAATTCATCAAACTCTGAGGAGCAGCATGAAATGGGGCTTA 1031
Db 1820 ATATAGTCTGTGAGGAGAAATTCATCAAACTCTGAGGAGCAGCATGAAATGGGGCTTA 1879
Qy 1032 CACAAAGATTCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1091
Db 1880 TACAGCAGATCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1939
Qy 1092 CGTGCAGGAAATGAAAGACCGAAACGCGGAAACCATCATGTCACACAGGACAGCTTAA 1151
Db 1940 TCGCGCGGAAATGAAAGACCGAAACGCGGAAACCATCATGTCACACAGGACAGCTTAA 1999
Qy 1152 TGCATTAACATGCTGAGAGTACGAAACCTCCGCAAGCTTTGGGCTTATGCGACAA 1211
Db 2000 TGCATTAACATGCTGAGAGTACGAAACCTCCGCAAGCTTTGGGCTTATGCGACAA 2059
Qy 1212 TATCAACATATGTCACAGTCTCTTCAAGGACGCAAAAGGATGAGAGGATCAAAAC 1271
Db 2060 TATTAACATATGTCACAGTCTCTTCAAGGACGCAAAAGGATGAGAGGATCAAAAC 2119
Qy 1272 GAATGTCATGACCGGAAAGCTTCATATACAGAAAGTCTTCTTGGACCGGATTC 1331
Db 2120 GGAAGTCATGACCGGAAAGCTTCATATACAGAAAGTCTTCTTGGACCGGATTC 2179
Qy 1332 CTCTCTGCGGCGAAACCATGACATTAATATATATATATATATATATATATATATATAT 1391
Db 2180 CTCTCTGCGGCGAAACCATGACATTAATATATATATATATATATATATATATATATAT 2239
Qy 1392 ACCGATGCTGTGGAACGTCGACAGCTTCGAGCATTTTGTGAAAAACAGAGGCAT 1451
Db 2240 TCCGATTTGTAGCAGGTATATGTGCAAAATTAAGGAGCATTTTGTGAAAAATGAGGGGT 2299
Qy 1452 CACACCAAGCTTCTCTATTAAGGCGGACATGATTCGCGTGCAGCTGACATCGGCT 1511
Db 2300 AACTCTTAAGCCTTCTCTATTAAGGCGGACATGATTCGCGTGCAGCTGACATCGGCT 2359
Qy 1512 TGGCTACCGGAAAGGTAAACAGGATGGGACAGATGACATTTGAATTAATCCCTGAACGT 1571
Db 2360 TGGCTTTTCCAAATGTAAACAGGATGGGACAGATGACATTTGAATTAATCCCTGAATGT 2419
Qy 1572 TGCCATATGTAAGCAGTCCAGTCTCTATCAACAGCCCAAAAGCGAGTACTGTTTAC 1631
Db 2420 CGCATTTGTGTAAGCAGTCCAGTCTCTATCAACAGCCCAAAAGCGAGTACTGTTTAC 2479
Qy 1632 TGCTACTGCGGCAAGCCTTTGAAATCTCCCTGTATGCTTGTATGCTCCCTGAGAC 1691
Db 2480 GGTCTCAAGCTGTGTAAGCCTTTGAAATCTCCCTGTATGCTTGTATGCTCCCTGAGAC 2539
Qy 1692 AACTGCTCGTAAAGCTGTGCAATGATGCACTTGTGCAATTAACGCTCAATATGCAAC 1751
Db 2540 GACGGCATCTCAATCTTTATGATGATGATTTGATGATGATGATGATGATGATGATGATG 2599
Qy 1752 ACAGTATGTAAGAAATGACCTTACTTGCCTATCAATGATTAATGAGATGCGCGCAATA 1811


```
XX JP2004008085-A.
PN
XX
XX 15-JUN-2004.
PD
XX
XX 06-JUN-2002; 2002JP-00165950.
PF
XX
XX 06-JUN-2002; 2002JP-00165950.
PR
XX
XX (KAOS ) KAO CORP.
PA
XX WPI; 2004-094297/10.
DR
XX
XX Novel mutant alkaline protease produced by substituting the amino acid
PT residues useful as washing agent.
PT
XX
XX Example 1; SEQ ID NO 8; 21pp; Japanese.
PS
XX
XX The invention comprises a mutant Bacillus alkaline protease which
CC contains substitutions at positions 163, 170 and 434. The mutant alkaline
CC protease is useful as a washing agent. The present DNA sequence
CC represents a mutagenic PCR primer for the Bacillus alkaline protease
CC gene.
XX
SQ Sequence 54 BP; 12 A; 17 C; 16 G; 3 T; 0 U; 6 Other;

Query Match      2.5%; Score 48; DB 12; Length 54;
Best Local Similarity 88.9%; Pred. No. 0.0007;
Matches 48; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1092 CGCTGCCGGGAAATGAAGACCGACGCGGAACATCATGTCAGCAGGACAGC 1145
DB 1 CGCTGCCGGGAAATNNNGACCGAACGCGGGAACATCCTNNNGACCAAGGACAGC 54

RESULT 15
AB082958/c
ID AB082958 standard; cDNA; 415 BP.
XX
XX AB082958;
AC
XX
XX 15-JAN-2003 (first entry)
DT
XX
XX Human lung specific nucleic acid (LSNA) SEQ ID NO:101.
DE
XX
XX Human; lung specific gene; lung specific nucleic acid; LSG; LSNA; LSP;
KM lung specific protein; lung cancer; vaccine; gene therapy; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200277236-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 26-OCT-2001; 2001WO-US050385.
PP
XX
XX 26-OCT-2000; 2000US-0243459P.
PR
XX
XX (DIAD-) DIADEXUS INC.
PA
XX
XX Rectipon H, Sun Y, Chen S, Liu C, Turner LR;
PI
XX WPI; 2003-018930/01.
DR
XX
XX New isolated nucleic acid molecule, useful for treating lung cancer, and
PT diagnosing or monitoring the presence of metastases of lung cancer in a
PT patient.
XX
XX Claim 1; Page 201; 262pp; English.
PS
XX
XX The present invention describes a isolated human lung specific nucleic
CC acid (LSNA) molecule comprising a sequence selected from AB082858 to
CC AB082993, which encodes a human lung specific protein (LSP) sequence
```

```
CC selected from ABP54165 to ABP54272. Also described: (1) a method for
CC determining the presence of a LSNA or LSP in a sample; (2) a vector
CC comprising the novel nucleic acid; (3) a host cell comprising the vector
CC of (2); (4) producing a polypeptide encoded by the novel nucleic acid;
CC (5) an isolated polypeptide produced by the method of (4); (6) an
CC antibody or its fragment that specifically binds to the polypeptide of
CC (5); (7) diagnosing and monitoring the presence of metastases of lung
CC cancer in a patient; (8) treating a patient with lung cancer; or (9) a
CC vaccine comprising the polypeptide of (5) or the novel nucleic acid. The
CC human LSNA and LSP sequences have cytosstatic activity, and can be used in
CC vaccines and gene therapy. The LSNA and LSP sequences can also be used
CC for treating lung cancer, and diagnosing or monitoring the presence of
CC metastases of lung cancer in a patient
XX
```

```
SQ Sequence 415 BP; 117 A; 64 C; 79 G; 155 T; 0 U; 0 Other;

Query Match      2.3%; Score 45; DB 8; Length 415;
Best Local Similarity 60.0%; Pred. No. 0.016;
Matches 75; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 486 TGTAAAGCAATGCGCTTGTATGAAAGCAAGAAAATTAAGAGGTGCAATTAAAGGCAAT 545
DB 255 TGAAGAAACATCCCAATTGATAGCAAGTAAAAATATCAATAGAGAAATTAATAGGCA 196

QY 546 CGAACAAATGCGCAATTCGCAATTAAGCAATGATGCTATATTAACGGCAAGCCTGA 605
DB 195 AGATTAAGCATCTCACAAATTATTAATCTAAGTTCTGCTGATTAATAATGCAAAATTTGA 136

QY 606 GTATA 610
DB 135 GAAAA 131
```

Search completed: April 8, 2006, 03:01:43
Job time : 2991.33 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 01:05:15 ; Search time 6945.92 Seconds

(without alignments)
12953.156 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923

Sequence: 1 atgaggaagaagaagaaggt.....cgttgcaattgtaataa 1923

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	44.8	2.3	1101	10	CNS0039G
2	43	2.2	486	5	BU682328 UI-CF-BC1
3	42.4	2.2	500	1	AI2659898
4	42.4	2.2	1201	10	CNS0102T
5	42	2.2	633	3	BU369190
6	42	2.2	631	3	BU393925
7	41.8	2.2	457	1	AI689929 rx31909.x
8	41.8	2.2	464	1	AA399583 rx31904.s
9	41.8	2.2	522	1	AI961026 w588d01.x
10	41.8	2.2	561	1	AM572909 hf17d03.x
11	41.8	2.2	628	6	CM440627 UI-H-BD0-
12	41.8	2.2	678	10	CM901981
13	41.8	2.2	685	10	CM536661
14	41.8	2.2	7301	4	HSMB05591
15	41.4	2.2	449	3	BU359939
16	41.4	2.2	997	10	CNS005TB
17	41.2	2.1	750	5	BQ743928
18	41.2	2.1	770	9	BZ069516
19	41.2	2.1	914	8	DN483960
20	41	2.1	440	1	AA865302
21	41	2.1	482	7	CR743947
22	41	2.1	486	5	BX112112

23	41	2.1	519	1	AI732955
24	41	2.1	531	1	AI762342
C 25	41	2.1	983	10	CL514519
26	40.8	2.1	625	7	CV250509
27	40.8	2.1	1101	10	CNS000CNT
28	40.6	2.1	538	1	AA037767
29	40.6	2.1	648	8	DR063487
C 30	40.6	2.1	1253	10	AG390521
C 31	40.4	2.1	341	7	CK479039
C 32	40.4	2.1	659	10	CM154791
C 33	40.4	2.1	1101	10	CNS00240
C 34	40.2	2.1	854	6	CB314560
35	40	2.1	433	1	AU300716
C 36	40	2.1	523	7	CM487402
C 37	40	2.1	575	10	CM856704
C 38	40	2.1	588	1	AU260495
C 39	40	2.1	764	7	CJ361543
C 40	40	2.1	766	7	CJ345334
C 41	40	2.1	944	7	CJ363875
C 42	39.8	2.1	484	3	BU334651
C 43	39.8	2.1	601	3	BU387574
C 44	39.8	2.1	674	5	BQ504480
C 45	39.8	2.1	828	1	AL706691

ALIGNMENTS

RESULT 1
CNS0039G/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TBT3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL063921
VERSION
AL063921.1 GI:4941778
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Bphyrroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
JOURNAL
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosawa and Aaron Kammoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain V2; on bw bp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_1b="RPCI-98"
/note="end : TBT3"

ORIGIN

```

/dev_stage="Adult and Fetal"
/1ab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-CF-EC1"
/note="Organ: Lung; Vector: pRTT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGCTTAC.
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371UI
and 380-383
TAG_LIB=UI-CF-EC1
TAG_SEQ=AAGTGCTTAC"

```

[illegible]

LOCUS	AI269898	500 bp	mRNA	linear	EST 28-JAN-1999
DEFINITION	q040401.x1 NCI_CGAP_k1d3 Homo sapiens cDNA clone IMAGE:1867969 3', mRNA sequence.				
ACCESSION	AI269898				
VERSION	AI269898.1	GI:3869065			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1 (bases 1 to 500)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D.				

Qy 1268 AACCGATGTCATGACCGGGAACGTTTCATACATGACGAAGATCTTCTTG 1321
 |||||
 Db 436 AACCTGATATAGTTGCACCTGCGATATATTTACATCGGCAAGATCAATATGCTG 489
 |||||

RESULT 6
 BU393925 641 bp mRNA linear EST 08-MAR-2002
 LOCUS BU393925 Dictyostelium discoideum cDNA library, SF Dictyostelium
 DEFINITION dictoideum cDNA clone ddb33105 5', mRNA sequence.
 ACCESSION BU393925
 VERSION BU393925 GI:19305011
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Eukaryota; Dictyostelidia; Dictyostelium.
 REFERENCE 1 (bases 1 to 641)
 AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-1, T.
 TITLE Full length cDNA of Dictyostelium discoideum at the slug stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadao Shin-1
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 source
 1..641
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddb33105"
 /sex="mat A"
 /dev_stage="slug stage"
 /clone_jib="Dictyostelium discoideum cDNA library, SF"

ORIGIN
 Query Match 2.2%; Score 42; DB 3; Length 641;
 Best Local Similarity 60.5%; Pred. No. 1.4;
 Matches 69; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1208 ACAATATCAACCATGTGACAGTTCTCTTCACTGACCGACCAAGATGACGATCA 1267
 |||||
 Db 169 ATTAATGAATAATATATGGATCATTTCTCATCAAGGCTCAACATATGTAATGCA 228
 |||||

Qy 1268 AACCGATGTCATGACCGGGAACGTTTCATACATGACGAAGATCTTCTTG 1321
 |||||
 Db 229 AACCTGATATAGTTGCACCTGCGATATATTTACATCGGCAAGATCAATATGCTG 282
 |||||

RESULT 7
 LOCUS A1689929 457 bp mRNA linear EST 16-DEC-1999
 DEFINITION tx31909.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2271232 3',
 mRNA sequence.
 ACCESSION A1689929
 VERSION A1689929 GI:4901223
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 457)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 www.bio.lnln.gov/btrp/image/image.html
 Insert Length: 577 Std Error: 0.00
 Seq primer: -40UP from G1bco.
 Location/Qualifiers
 1..457
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2271232"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /clone_jib="NCI CGAP Lu24"
 /note="Organ: lung; Vector: pT73D-Pac (pharmacia) with a
 modified polylinker; Plasmid DNA from the normalized
 library NCI-CCAP Lu5 was prepared, and 86 clones were
 made in vitro. Following HAP purification, this DNA was
 used as tracer in a subtractive hybridization reaction.
 The driver was PCR-amplified cDNAs from a pool of 5,000
 clones made from the same library (cloneIDs
 141920-1417991 and 1520904-1522439). Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 2.2%; Score 41.8; DB 1; Length 457;
 Best Local Similarity 57.1%; Pred. No. 1.4;
 Matches 76; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 174 TGCTGCTTTCTGTGTGAATGAAATGTAATCTCCCAAGGTTGCAAGAAAGCT 233
 |||||
 Db 325 TGCACTTATGTGTGATGAAAACAAGCTAATGAAACACTGATCTGATTTAAAGTTT 384
 |||||

Qy 234 TGAACAGTCCCGCAATATTAATCTTCATTTATTCATTTCAATGACCAATTTTGA 293
 |||||
 Db 385 TCTACCTGTATACACTCTACTACATCTCCATACCAACCAACTTTGGTCCATTAACA 444
 |||||

Qy 294 AGAAACAAACAG 306
 |||||
 Db 445 TGAGCAAAACAG 457
 |||||

RESULT 8
 LOCUS AA399583 464 bp mRNA linear EST 16-MAY-1997
 DEFINITION zt93b04.e1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:729871
 3', mRNA sequence.
 ACCESSION AA399583
 VERSION AA399583 GI:2052587
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 464)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Joet, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, K., Stepec, M., Tan, F., Theising, B.,
 White, Y., Wyllie, T., Waterston, R. and Wilson, R.
 TITLE WashU-Merck EST Project 1997
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson R.
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNLN; contact the

IMaGS Consortium (info@image.llnl.gov) for further information
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 451.

FEATURES

Bourse

```
1. 464
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="GB:592640"
/db_xref="Exon:9606"
/clone="IMAGE:729871"
/sex="male"
/sex_bac="none"
```

ORIGIN

Query Match	2.2%	Score 41.8;	DB 1;	Length 464;
Best Local Similarity	57.1%;	Pred. No. 1.4;		
Matches 76;	Conservative 0;	Mismatches 57;	Indels 0;	Gaps 0;

QY	17	TGCTGCTTTTCTGCTGGAATCTGAANAATGGAACCTCCCAAAAGTTTGCAGAAAGAGCT	233
Db	319	TGCACTTATGTGTATGAAAACAAAGCTAACTGAAACCTGCAGTGTATTTAAAAAGTTT	378
QY	234	TGAACAGTCCCGGCAATATATTAATCTCATTTATTCNAATTCAATGACCAATTTTGA	292
Db	379	TCTACCTGTACACACTACTACTACCTCATTCAAAACACAACTTTGGTCCATPAAACA	438
QY	294	AGAAACAAAACAG	306
Db	439	TGAGACAAAACAG	451

LOCUS	A1961026	522 bp	mRNA	linear	EST-08-MAR-2000
DEFINITION	wg98d01.x1 NCI_CGAP_GCC Homo sapiens cDNA clone IMAGE:2475457 3' , mRNA sequence.				
ACCESSION	A1961026				
VERSION	A1961026.1	GI:5753807			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 522) NCI-CGAP, http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)				
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: W. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ LMU at: www-bio.ljml.gov/bbrp/image.html Insert Length: 2038				
COMMENT	Std Error: 0.00				

Seq primer: -40UP from Gibco
High quality sequence stop: 462.

FEATURES

Source

```
/organism="Homo sapiens"  
/mol_type="RNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:2475457"  
/tissue_type="pooled germ cell tumors"  
/lab_host="nufr"
```

ORIGIN

Query Match	2.2%	Score 41.8;	DB 1;	Length 522;
Best Local Similarity	57.1%;	Pred. No. 1.5;		
Matches 76;	Conservative 0;	Mismatches 57;	Indels 0;	Gaps 0

OY	174	TGCGCTTTCTGGTGGATCTGAAAATGTAAATCTCCAAAGGTTTGAGAAGAGCT	233
Db	219	TGCGATTATGTGTATGAAAACAGCTAACTGAAACCTGCATGCTGATTTTAAAGTTT	278
OY	234	TGAACAGTCCGGCAATTAATTAATCTCATATTATTCAAATTCAATGACCAATTTTAGA	293
Db	279	TCTACTGTATACACTACTACTACTCTCCATCAACACCAAACTTTGGTCAATTAACCA	338
OY	294	AGAAACAAAACAG	306
Db	339	TGAGCAAAAACAG	351

```

RESULT 10
AM572909
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
ORGANISM
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 561)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LML, contact the
IMAG Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 449.
Location/Qualifiers
1..561
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2932133"
/lab_host="DH10B"
/clone_1lb="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pVT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;"

```

Equal amounts of plasmid DNA from three normalised libraries (fetal lung NbH1.9w, testis NHT, and B-cell NCI-GARP GCBI) were mixed, and six circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 726408-G.B. clones 297480-102087, 682632-687239, 726408-728711, and 729096-731393. Subtraction by Bento Soares and M. Fatima Bonaldo "

ORIGIN

Query Match	2.2%	Score 41.8	DB 1	Length 561
Best Local Similarity	57.1%	Pred. No. 1.5	57	Indels 0
Matches	76	Conservative	0	Mismatches 0
				Gaps 0
QY	174	TGCTGCTTTTCGTGATCTGAAAATGTGAAACTCCCAAAAGTTTGACAGAAAGCT	233	
Db	319	TGCAGTTATGTTGATGAAACAACTACTAAACCTGCATGTGATTTTAAAGTTT	378	
QY	234	TGAACAGTCCCGCAATTAATTAATTCATATTATTCATTCATGACCAATTTTGA	293	
Db	379	TCTACGTGTACACTACTACTACACTCATACAAACCAACTTTGGTCCATTAACA	438	
QY	294	AGAAACAAAACAG	306	
Db	439	TGAGACAAAACAG	451	

RESULT 11	CA440627	LOCUS	DEFINITION
CA440627	628 bp	mRNA	linear
UI-H-ED0-axd-1-08-0-UI.s1	NCI CGAP	ED0 Homo sapiens	cdna clone
UI-H-ED0-axd-1-08-0-UI.3'	mRNA	sequence.	

ACCESSION	CA440627	
VERSION	CA440627.1	GI:24805047
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap .
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE	Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgapbe-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
PolyA+yes.

FEATURES	Location/Qualifiers
source	1. .628

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-ED0-axd-1-08-0-UI"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED0"
/notes="Organ: Left Pubic Bone; Vector: pYT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Bscr I;
Site 2: Not I; NCI CGAP ED0 is a cDNA library containing
the following tissue(s): Chondrosarcoma cell line C5. The
library was constructed according to Bonaldo, Lennon and
Scorel, Genome Research, 6:791-806, 1996. First strand

```

cDNA synthesis was primed with an oligo-dT primer
 containing a Not I site. Double stranded cDNA was ligated
 to an EcoR I adaptor, digested with Not I, and cloned
 directionally into pT73-Pac vector. The oligonucleotide
 used to prime the synthesis of first-strand cDNA contains
 a library tag sequence that is located between the Not I
 site and the (dT)18 tail. The sequence tag for this
 library is GCTCAAGGCT.
 TAG_T15SUB=chondrosarcoma
 TAG_L1B=U1-H-BD0
 TAG_SEQ=CGTCAGGCGCT"

ORIGIN

Query Match	2.2%	Score 41.8;	DB 6;	Length 628;
Best Local Similarity	57.1%	Pred. No. 1.6;		
Matches	76;	Conservative	0;	Mismatches 57; Indels 0; Gaps 0
QY	174	TGCTGCTTTCTGTGTGATCTGAAATGTAATCCCAAAAGTTTGCAGAGAAGCT	233	
Db	334	TGCAGTTATGGTGAATGAAACAGCTAACTAAACACTGCACTGTCTATTTTAAAGTTT	393	
QY	234	TGAAACAGTCCCGCAATTAATAATTCATATTATTCAAATTCAGAGCAATTTTAG	293	
Db	394	TCTACTCTGTACACTACTACTACTACCTCATCAAAACAACTTTGTCCATATAAC	453	
QY	294	AGAAACAAACG	306	
Db	454	TGAGCAAAACG	466	

RESULT 12				
CM901981/c				
LOCUS	CM901981	678 bp	DNA	linear GSS 08-DEC-200
DEFINITION	RPCT42_135B1.TJ RPCT-42 Bos taurus genomic clone RPCT42_135B1,			
	genomic survey sequence.			

ACCESSION	CM901981
VERSION	CM901981.1
KEYWORDS	GI:56444251
SOURCE	GSS.
ORGANISM	Bos taurus (cow)
	Bos taurus

REFERENCE	
AUTHORS	Larkin,D.M., Donthu,K., Lebduc,R., Ryan,K., Liu,L., de Jong,P.J. and Lewin,H.A.
TITLE	End sequencing of Holstein BAC library RPCI-42
JOURNAL	Unpublished (2004)
COMMENT	Other GSSs: RPCI42 13961.TV Pecora: Bovidae; Bovinae; Bos. 1 (bases 1 to 678) Laurasiatheria; Cetartiodactyla; Ruminantia;

Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library RPC1-42
(<http://bncmac.chori.org/bncvln42.htm>) For BAC library

availability; please contact Pieter de Jong (pjd@original.chi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm). Funding was provided in part by grant no. AG0004-34480-14417 from USDA-CERES (Livestock Genome Sequencing Initiative) and AG05-5438-2-313 from USDA-ARS Plate: 139 row: B column: 1
Class: BAC ends
Seq primer: SP6

```

FEATURES
  source
    Location/Qualifiers
      1..678
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /strain="Holstein"
        /db_xref="taxon:9913"

```

/clone="RPC142_139B1"
 /sex="male"
 /cell_type="blood"
 /clone_id="RPC1-42"
 /note="Vector: DBACE3.6; Site 1: EcorI; Site 2: EcorI;
 RPC1-42 Bovine BAC library (Male) produced by Pieter de
 Jong"

ORIGIN

Query Match 2.2%; Score 41.8; DB 10; Length 678;
 Best Local Similarity 56.0%; Pred. No. 1.6;
 Matches 79; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 197 AAATGGAACCTCCAAAGTTTGAGAGAAAGCTTGAACAGTCCCGCAATATA 256
 |||||
 DB 491 AAAAGAAAACTATTAAAGAAATGAGACATTCAGAAACCTCCAGACAAATTTA 432
 |||||
 QY 257 AACTCATATTATTCATTCATATGACCAATTTTGAAGAAACAGACAGCTGGAA 316
 |||||
 DB 431 AAGCCCAACGTTCAATCATAGAGTCCCAAGAGACAGAAAGAACCATGAGA 372
 |||||
 QY 317 AAACAGGGGCAAGATTCTCG 337
 |||||
 DB 371 AATACTTGAGAGATATATG 351
 |||||

RESULT 13
 CWS36661/c 685 bp DNA linear GSS 08-OCT-2004
 LOCUS OP_Ba0033K11.r OP_Ba Oryza punctata genomic clone OP_Ba0033K11
 DEFINITION 3', genomic survey sequence.

ACCESSION CWS36661.1 GI:54010883
 VERSION CWS36661.1
 KEYWORDS GSS.
 SOURCE Oryza punctata
 ORGANISM Oryza punctata
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 685)
 Samiguel, P., Westerman, R., Kim, H., Yu, Y., Wisotski, M., Yost, D.,
 Stum, D., Rao, K., Luo, M., Jettly, R., Kudrna, D., Miller, C.,
 Hatfield, J., Soderlund, C., Wang, R. and Jackson, S. A.
 Hatfield, J., Soderlund, C., Wang, R. and Jackson, S. A.
 OMAF Project - Purdue University
 Unpublished (2004)
 CONTACT: Scott A. Jackson
 JACKSON Laboratory
 Purdue University
 915 W. State St., West Lafayette, IN 47907, USA
 Tel: 7854963621
 Fax: 7654967255
 Email: sjackson@purdue.edu

TITLE
 JOURNAL
 COMMENT
 PCR Primers
 FORWARD: TAA TAC GAC TCA CTA TAG GG
 BACKWARD: CAC TCA TTA GGC ACC CCA
 Insert length: 161000 Std Error: 0.00
 Place: 0033 row: K column: 11
 Seq primer: CAC TCA TTA GGC ACC CCA
 Class: BAC ends.

FEATURES
 source
 1..685
 Location/Qualifiers
 /organism="Oryza punctata"
 /mol_type="genomic DNA"
 /db_xref="taxon:4537"
 /clone="OP_Ba0033K11"
 /issue_type="young leaves"
 /lab_host="DH10B-T1 phage resistant"
 /clone_id="OP_Ba"
 /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 2.2%; Score 41.8; DB 10; Length 685;
 Best Local Similarity 49.3%; Pred. No. 1.6;
 Matches 109; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 156 GCAGGGGACAGCTGGTGCTGCTCTTTCGGTGGAAATCGAAATGAAATCTCCAA 215
 |||||
 DB 615 GCAGATGTAAAGGGGCTCTTGTATTAACTTCACACAGCAATGTCAAGTGGCAT 556
 |||||
 QY 216 AGTTTGACAGAGACCTTGAACAGTCCCGCAATTAATTAATCTCATTTTCAAT 275
 |||||
 DB 555 CAGCATTTAGAAACAGTGCTTCAACTCAGACCATCTTGCAAGTCACTACCTGT 496
 |||||
 QY 276 CAATGACCAATTTAGAAAGAAACAAACAGCAGCTGGAAAAACAGGGGCAAGTTCT 335
 |||||
 DB 495 ACAATTTCAACCTTTTGAAGAAAGAAATCAGATTTTGAAGAAATCATTTGCAACATAA 436
 |||||
 QY 336 CCACTACATACCTGATTTATCTTCACTTGTCCAGTATGAG 376
 |||||
 DB 435 CTTATATCAACCAATCATGCAAGAAATTAATCAATATGTAAG 395
 |||||

RESULT 14
 HSM805591/c 7301 bp mRNA linear HTC 22-SEP-2004
 LOCUS HSM805591
 DEFINITION Homo sapiens mRNA; cDNA DKFZp434J1323 (from clone DKFZp434J1323).
 ACCESSION AL834204
 VERSION AL834204.1 GI:21739712
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE 1 (bases 1 to 7301)
 Blum, H., Bauerbach, S., Mewes, H. W., Weil, B., Amid, C., Oseanger, A.,
 Fobo, G., Han, W. and Wiemann, S.
 The German cDNA Consortium
 Direct Submission
 Submitted (22-SEP-2004) MIBS, Ingolstaedter Landstr.1, D-85764
 Neuberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by LMU (Ludwig Maximilians University, Munich/Germany)
 within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp434J1323) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434J1323
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.

TITLE
 JOURNAL
 COMMENT
 Neuberger, GERMANY
 Submitted (22-SEP-2004) MIBS, Ingolstaedter Landstr.1, D-85764
 Neuberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by LMU (Ludwig Maximilians University, Munich/Germany)
 within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp434J1323) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434J1323
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.

FEATURES
 source
 1..7301
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp434J1323"
 /db_xref="taxon:9606"
 /clone="DKFZp434J1323"
 /issue_type="testis"
 /clone_id="434 (synonym: hbea3). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /note="ankyrin repeat-containing protein 12"

gene
 CDS
 1..7301
 /gene="DKFZp434J1323"
 /join(193..3069,3069..6311)
 /gene="DKFZp434J1323"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAH56382.1"
 /db_xref="GI:52545871"
 /translation="MPKSGFTYQIENSNSDSNMYEKRYGRKSKDIASYSKTPKIE

RSDVSEKEMKESGSMKRLPFTISPSRNERBDSDEKEGPEKKTKKAGNKKSTPVSI
 LRGVLSRKMOMALMOMTARDNSPDSTPNHPSOTTPAKOTKTPSSSQKXKVNRE
 RGETPLHMAALIKGVYKQYKELISLGNVNVDPACMTPLHRCANVTYDVAKILLAG
 ADVNTQGLDDDTPLHDSASGHRDIKLLRRGNPFOANKGEPVDAETBELLL
 LKREVPGLSDDESYTDSBAQSVNPSVDENISDETKDSLCEKQILPSTPLPSA
 LDEYFPLKDDDEYKMDIDRAILKREKQENPDAEKTLPAKOEKAFYPSFPGK
 OQPSVLYSTDSDEBALONKSTISCSVIPYTNPSMDQTKGYVSEHOKQYK
 RKLKONKKNENKQELKOBKGEKENTRTNLTNTGLDCSEKTEREGRNPKSSPKDDT
 SLHLPHISTGKSPKSCGLSEKOSTPLKQEHYKTLSPSSSEMSLQPLVRYDNTSE
 FLPESSSVKSCNKKKSRHQDPLHPEGKSNVAKIDEBHSPTFENSDDLKQMDXG
 KTLKQKQLKHKREKSKHKELEGKEKXTDQSAKELQSVPEPRBEWKENPLKQXG
 TEDLPLNMEHSLTLEKSKLEKNIKQDSTGEKYSKERNFSEKEDKIKESSEKFR
 BEKIKOLKEBENITPDKOSSEPTSGMSAIEESIGLHVEKIDIEKOKHKESEK
 PEKROIKERKOEKREKTEREKKIKHHEKSKDKLDSRVDKIKEDKLYSHHT
 KCHKEGSKRTAALIKTTDDRKSRRDRKDKKREKREKRLAESKEXHLEKQKQ
 SDNSBYSKSEKKNREKQELDKESKSDKESINTNSKHIOEKSSSVLDGNKQHE
 KPLSLKERTKQEPKLPDQEKDKDKIDRYEKDKDKDKQINSLLKLSKADPK
 PXSASKQDTRKEXKRLVNDLMDQTSFERMLSLKOLEIQMHKXKXKIKOKEKERLR
 NNNCLBKIKDKERTKPTESKNEKELTFSKSEVTDATYKOPKPDVNSKOSVDT
 KNMTLGKSSPVSDNSLNRSPRSNKKQLSRSVSMISVASSSESCHTTTPPBPV
 EYSDPMLSESSQMSFOSPFYSLAKSPALHERLDSLADIPERIKPPYANLSTSH
 LSSSEVEDKLLISGRPTIEVRCSMPVCEHTKQFOTISESNQSLTVPEDTS
 PSPKEVPSNVEPRDLNVSNTSHSFATSPGASNKYSADRNLIKOTAPVNTWDS
 PVHLEPSSQVGVIOKNSWEMPYDRLETLSTRDFICPNSNIPQESLSQSFENSEKVL
 KENADPLSLRQTELEPNSCAODPASFMPQCPSPPOSLSDAESISKMSLSVYANQ
 BEGILQONKNAVQIISALDPTNESTKQDENTPFVLDVOKTDAFVYVSDSTIOEASP
 PEKATYTLVPLSEKQDNGSDASTOLANTHYAPSKLYKSSGHEVENSTTPQVISHK
 EKLESILVITHLSRCDSDLCENAMGPKCNLMBOPKACPBSEKCLISIDRESQOSI
 LSLLENHSQOSTOPMHTKQGLVKEVLENAEDDTENQIPQMRNRNANTYANOSKO
 ILASCTLSSEKDSSESSPRGIRLTEDDDPQIHPRKRVSVSPVQVSPSLQAKS
 KTOQSLAIVDSLKLEIDLOPYGSEERANPFYELHARKIEERKLCISVTPAPQYD
 EYVTNGSLVLDGNPLSKICPTTPPRLPELKEPQCEVVMKRLQHSIEREK
 LIVSNEQVLRVYRAARTLANOTLPKASCYLLDAEYVNVPLDOSDSKTSVDRP
 NARQPMSTLQVDDVDPDKITCLMLRQCHEAALAAVQRLQELQELDPATYKSTIS
 IYEQFPYVPLVNDPELTPPI"

ORIGIN

Query Match 2.2%; Score 41.8; DB 4; Length 7301;

Best Local Similarity 57.1%; Pred. No. 3.7; Mismatches 0; Gaps 0;

Matches 76; Conservative 0; Indels 57; Indels 0; Gaps 0;
 QY 174 TGCCTCTTTCTGTGTAATCTGAAATGTGAAACTCCCAAAAGTTTGACAGAGAGCT 233
 |||||
 DB 7058 TGCAGTTATGTGATGAAAAGAGCTAAGCACTGAAACCTGCAATGCTGATTTAAAGTT 6999
 |||||
 QY 234 TGAACAGTCCCGGCAATTAATCTCATTTATCCATTCATTCAGCAATTTTACA 293
 |||||
 DB 6998 TCTACCTGTACACACTACTACTACCTCATCAAAACCAACTTGTGTCATTAACA 6939
 |||||
 QY 294 AGAACAACAACAG 306
 |||||
 DB 6938 TGAACAACAACAG 6926

RESULT 15

BJ359939

LOCUS BJ359939 449 bp mRNA linear EST 07-MAR-2002

DEFINITION BJ359939 Dictyostellium discoidium cDNA library, CF Dictyostellium

discoidium cDNA clone ddc3k09 5', mRNA sequence.

ACCESSION BJ359939.1 GI:19259534

VERSION BJ359939.1 GI:19259534

KEYWORDS EST.

SOURCE Dictyostellium discoidium

ORGANISM Dictyostellium discoidium

REFERENCE 1 (bases 1 to 449)

AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE Full length cDNA of Dictyostellium discoidium at the culmination

stage

JOURNAL Unpublished (2002)

COMMENT Contact: Tadaeu Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 location/Qualifiers

FEATURES

source

1..449
 /organism="Dictyostellium discoidium"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="ddc3k09"
 /sex="mat A"
 /dev_stage="Culmination stage"
 /clone_lib="Dictyostellium discoidium cDNA library, CF"

ORIGIN

Query Match 2.2%; Score 41.4; DB 3; Length 449;

Best Local Similarity 60.0%; Pred. No. 1.9; Mismatches 46; Indels 0; Gaps 0;

Matches 69; Conservative 0; Indels 46; Indels 0; Gaps 0;

QY 1208 ACAATATCAACATGTGACACAGTTCTTCACTGACGACCAAGATGACAGATCA 1267
 |||||
 DB 89 ACAATGAGAAATATATTTGTTCAATTCATCAAAAGTCCAAACACATGATGATGAATGA 148
 |||||
 QY 1268 AACCGATGTCATGACACCGGGAAGCTTCATCTATCAAGCAAGATCTTCTTGC 1322
 |||||
 DB 149 AACCTGATTTAGTCCCTGCTGTGATATATTTACATGGCAAGATCTAATGTGC 203

Search completed: April 8, 2006, 08:09:29

Job time : 6949.92 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 00:12:00 ; Search time 589.96 Seconds
(without alignments)
5794.037 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923
Sequence: 1 atgagaagaagaagaagaagt.....cgttcgcaatcgtgaataa 1923

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/6 COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1923	100.0	1923	3	US-09-509-814A-5
2	1923	100.0	1923	3	US-09-920-954-5
3	1919.8	99.8	1923	3	US-09-509-814A-7
4	1919.8	99.8	1923	3	US-09-920-954-7
5	1477.2	76.8	1920	3	US-09-509-814A-3
6	1477.2	76.8	1920	3	US-09-920-954-3
7	1166.4	60.7	3003	2	US-08-873-479-41
8	54	2.8	7218	2	US-08-232-463-14
9	42.4	2.2	834	3	US-09-621-976-2574
10	41.8	2.2	5053	3	US-09-620-312D-590
11	41.4	2.2	108341	3	US-09-949-016-16621
12	39.8	2.1	90724	3	US-09-949-016-16601
13	38.8	2.0	1141	3	US-09-806-708B-22
14	37.6	2.0	2835	2	US-08-750-532-2
15	37.6	2.0	4765	2	US-08-750-532-8
16	37.6	2.0	4765	3	US-08-894-818B-7
17	37.6	2.0	4765	3	US-09-445-472-5
18	37.6	2.0	4765	3	US-10-090-624-5
19	37.6	2.0	4765	3	US-09-841-553-7
20	37	1.9	9623	3	US-08-956-171B-166
21	36.6	1.9	9623	3	US-08-781-986A-166
22	36.6	1.9	2001	3	US-09-949-016-2801
23	36.6	1.9	34068	3	US-09-949-016-15489
24	36.6	1.9	51711	3	US-09-949-016-12559

25	36.4	1.9	12167	3	US-09-949-016-15180	Sequence 15180, A
26	36.4	1.9	39480	3	US-09-949-016-13634	Sequence 13634, A
27	36.4	1.9	385136	3	US-09-949-016-16073	Sequence 16073, A
28	36.2	1.9	474	3	US-09-621-976-18033	Sequence 18033, A
29	36	1.9	699	3	US-09-328-352-2529	Sequence 2529, Ap
30	36	1.9	2262	3	US-09-252-991A-1416	Sequence 1416, Ap
31	36	1.9	2391	3	US-08-750-532-1192	Sequence 1192, Ap
32	35.6	1.9	564	2	US-08-750-532-11	Sequence 11, Appl
33	35.6	1.9	564	3	US-08-894-818B-14	Sequence 14, Appl
34	35.6	1.9	564	3	US-08-841-553-14	Sequence 14, Appl
35	35.6	1.9	898	2	US-08-750-532-7	Sequence 7, Appl
36	35.6	1.9	1859	3	US-08-894-818B-15	Sequence 15, Appl
37	35.6	1.9	1859	3	US-09-841-553-15	Sequence 15, Appl
38	35.6	1.9	1977	3	US-08-894-818B-2	Sequence 2, Appl
39	35.6	1.9	1977	3	US-08-894-818B-6	Sequence 6, Appl
40	35.6	1.9	1977	3	US-09-445-472-11	Sequence 11, Appl
41	35.6	1.9	1977	3	US-10-090-624-11	Sequence 11, Appl
42	35.6	1.9	1977	3	US-09-841-553-2	Sequence 2, Appl
43	35.6	1.9	1977	3	US-09-841-553-6	Sequence 6, Appl
44	35.4	1.8	601	3	US-09-949-016-44315	Sequence 44315, A
45	35.4	1.8	735	3	US-09-328-352-415	Sequence 415, App

ALIGNMENTS

RESULT 1						
US-09-509-814A-5						
Sequence 5, Application US/09509814A						
Patent No. 6376227						
GENERAL INFORMATION:						
APPLICANT: TAKATA, MIKIO						
APPLICANT: OKUDA, MITSUYOSHI						
APPLICANT: SAKI, KATSUHIISA						
APPLICANT: KUBOTA, HIROMI						
APPLICANT: HITOMI, JUN						
APPLICANT: KAGEYAMA, YASUSHI						
APPLICANT: SHIKATA, SHITSUM						
APPLICANT: NOMURA, MASAFUMI						
TITLE OF INVENTION: ALKALINE PROTEASE						
FILE REFERENCE: 0327-0832-0PCT						
CURRENT APPLICATION NUMBER: US/09/509,814A						
CURRENT FILING DATE: 2000-04-06						
PRIOR APPLICATION NUMBER: PCT/JP98/04528						
PRIOR FILING DATE: 1998-10-07						
PRIOR APPLICATION NUMBER: JP 9-274570						
PRIOR FILING DATE: 1997-06-08						
NUMBER OF SEQ ID NOS: 24						
SOFTWARE: PatentIn version 3.0						
SEQ ID NO 5						
LENGTH: 1923						
TYPE: DNA						
ORGANISM: Bacillus sp.						
FEATURE:						
NAME/KEY: CDS						
LOCATION: (1)..(1923)						
US-09-509-814A-5						
Query Match						
Best Local Similarity 100.0%; Score 1923; DB 3; Length 1923;						
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	ATGAGAAAGAAAAAGGTTTATCTGTTTATCTGTTTATCTGAGCTGACGATTTGTCAGT	60			
DB	1	ATGAGAAAGAAAAAGGTTTATCTGTTTATCTGTTTATCTGAGCTGACGATTTGTCAGT	60			
QY	61	GTTGGCTTAAGTAATCCATCTGACAGTGTGCAAGAAATTTGATCTGATTTCAAGGA	120			
DB	61	GTTGGCTTAAGTAATCCATCTGACAGTGTGCAAGAAATTTGATCTGATTTCAAGGA	120			
QY	121	ATTGAGACAAGCTGATGCTAAAGTTTCTCAAGCAGGAGGAGACTGCTGCTGCT	180			
DB	121	ATTGAGACAAGCTGATGCTAAAGTTTCTCAAGCAGGAGGAGACTGCTGCTGCT	180			

QY 181 TTCTGTGGGAATCTGAAAAATGAAACTCCAAAAGTTTGGAGAAGAAGCTTGAACA 240
DB 181 TTTCTGTGGAACTGAAAAATGAAACTCCAAAAGTTTGGAGAAGAAGCTTGAACA 240
QY 241 GTCCCGCAATATATTAACCTCCATATTATCAATTCAATGAACCAATTTTAAAGAAACA 300
DB 241 GTCCCGCAATATATTAACCTCCATATTATCAATTCAATGAACCAATTTTAAAGAAACA 300
QY 301 AAACAGCAGCTGAAAAAACAAGGAGCAAGATTCTGACTACATACCTGATTAAGTTTAC 360
DB 301 AAACAGCAGCTGAAAAAACAAGGAGCAAGATTCTGACTACATACCTGATTAAGTTTAC 360
QY 361 ATTGTGATATAGAGGCGCATGTTAAGTCAACAACAAGCAACATTGAGCAGCTGGAATCC 420
DB 361 ATTGTGATATAGAGGCGCATGTTAAGTCAACAACAAGCAACATTGAGCAGCTGGAATCC 420
QY 421 GTGAGACCTTATTTTCCGATATACAGAAATAGATCCCACTTTTCAAAAAGGGGATCA 480
DB 421 GTGAGACCTTATTTTCCGATATACAGAAATAGATCCCACTTTTCAAAAAGGGGATCA 480
QY 481 GAGCTTGTAAAGCAGTGGGCGCTGTATACAAAGCAAGAAAAATTAAGAGTGCATTTAGA 540
DB 481 GAGCTTGTAAAGCAGTGGGCGCTGTATACAAAGCAAGAAAAATTAAGAGTGCATTTAGA 540
QY 541 GGCATCGAACCAATCGCAATTTCCGATATAGCAATGATGTCTATATATTAACGGCAAG 600
DB 541 GGCATCGAACCAATCGCAATTTCCGATATAGCAATGATGTCTATATATTAACGGCAAG 600
QY 601 CTTGAGTATAGGTATGATGATGATGTTGGCGCTGGAATTTCAAAAAGGAGTGGCTGAG 660
DB 601 CTTGAGTATAGGTATGATGATGATGTTGGCGCTGGAATTTCAAAAAGGAGTGGCTGAG 660
QY 661 AGCAGTACGGGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 AGCAGTACGGGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACAGGTGCGAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCGAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTGGAGCGGACCAATATATGCAATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGAGCGGACCAATATATGCAATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATATTAGAAACGGCTCCACTAATTAAGGAATGGCGCTCAGGCGAATCTAGCTTC 900
DB 841 TCCGATATTAGAAACGGCTCCACTAATTAAGGAATGGCGCTCAGGCGAATCTAGCTTC 900
QY 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGGAGACTTACCTTCAATCTGCAACCTTA 960
DB 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGGAGACTTACCTTCAATCTGCAACCTTA 960
QY 961 TTCAGCCAAAGCATACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 TTCAGCCAAAGCATACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 AATGGGCGCTTACCAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGCGCTTACCAACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACAGTCTTTTCTGCTGCGGGAATGAGAACCGGACCGGAAACCATGAGCAACAGGC 1140
DB 1081 ACAGTCTTTTCTGCTGCGGGAATGAGAACCGGACCGGAAACCATGAGCAACAGGC 1140
QY 1141 ACAGTCTTTTCTGCTGCGGGAATGAGAACCGGACCGGAAACCATGAGCAACAGGC 1200
DB 1141 ACAGTCTTTTCTGCTGCGGGAATGAGAACCGGACCGGAAACCATGAGCAACAGGC 1200
QY 1201 TATGCGGACCAATATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 TATGCGGACCAATATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1260

QY 1261 CGGATCAAAACCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 CGGATCAAAACCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GCACCGGATTCCTCTCTGCGGCGAACCATGACAGTAATATGATGATGATGATGATGAT 1380
DB 1321 GCACCGGATTCCTCTCTGCGGCGAACCATGACAGTAATATGATGATGATGATGATGAT 1380
QY 1381 TCCATGCTTACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 TCCATGCTTACACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 AACAGAGCATCAACCAAGCCTTCTATTAATAAGGCGACATGATGATGATGATGATGAT 1500
DB 1441 AACAGAGCATCAACCAAGCCTTCTATTAATAAGGCGACATGATGATGATGATGATGAT 1500
QY 1501 GACATCGGCTTTGGCTACCGAACGTTAACCAAGATGATGATGATGATGATGATGATGAT 1560
DB 1501 GACATCGGCTTTGGCTACCGAACGTTAACCAAGATGATGATGATGATGATGATGATGAT 1560
QY 1561 TCCCTGAAAGTTCCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1561 TCCCTGAAAGTTCCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 TACTGTTTACTGCTACCTGCGGCAAGCTTTGAAATCTCCCTGATGATGATGATGATGAT 1680
DB 1621 TACTGTTTACTGCTACCTGCGGCAAGCTTTGAAATCTCCCTGATGATGATGATGATGAT 1680
QY 1681 CTTGCGAGCAACATGCTCTCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGAT 1740
DB 1681 CTTGCGAGCAACATGCTCTCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGAT 1740
QY 1741 CCAATGCGACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 CCAATGCGACACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GCGCGCATTAACGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 GCGCGCATTAACGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTATGATGATGATGATGATGATGATGATGATGATGAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 2
US-09-920-954-5
Sequence 5, Application US/09920954
Patent No. 6739228
GENERAL INFORMATION: MIKIO
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKELI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/920, 954
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/509, 814
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24

RESULT 3
US-09-509-814A-7
Sequence 7, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIKA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-09-509-814A-7

Query Match 99.8%; Score 1919.8; DB 3; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCTGAGCGGCAATTTTGTGAGCT 60
DB 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCTGAGCGGCAATTTTGTGAGCT 60
QY 61 GTTGGTTAAGTATCACTGTCAGAGTGTGCAAGAAATTTGATCTGATTTCAAGAA 120
DB 61 GTTGGTTAAGTATCACTGTCAGAGTGTGCAAGAAATTTGATCTGATTTCAAGAA 120
QY 121 ATTGAGCAACAAGTCTGATGCTTAAAGTTTCTTCAAGCGGCGCAGACTGTGCTGCT 180
DB 121 ATTGAGCAACAAGTCTGATGCTTAAAGTTTCTTCAAGCGGCGCAGACTGTGCTGCT 180
QY 181 TTTCTGTGGAATCTGAAATGTGAATCTCCCAAAAGGTTTGCAAGAGAGCTTGAACA 240
DB 181 TTTCTGTGGAATCTGAAATGTGAATCTCCCAAAAGGTTTGCAAGAGAGCTTGAACA 240
QY 241 GTCCCGGCAAAATATTAATCTCCATTTATTCATTCATTCGAAATTTAGAGAAACA 300
DB 241 GTCCCGGCAAAATATTAATCTCCATTTATTCATTCATTCGAAATTTAGAGAAACA 300
QY 301 AAACAGCAGCTGGAAGAAAGAGGCGCAAGATTTCTGCACTATACCTGATTAAGCTTAC 360
DB 301 AAACAGCAGCTGGAAGAAAGAGGCGCAAGATTTCTGCACTATACCTGATTAAGCTTAC 360
QY 361 ATTGTGAGATATGAGGCGGATGTTAAGTCAAGCAACAGCAACATTTAGAGCAATTC 420
DB 361 ATTGTGAGATATGAGGCGGATGTTAAGTCAAGCAACAGCAACATTTAGAGCAATTC 420
QY 421 GTGAGAGCTTATTTGCGGATATACGAATAGATCCCGAGCTTTTCAAAAAGGCGCATCA 480
DB 421 GTGAGAGCTTATTTGCGGATATACGAATAGATCCCGAGCTTTTCAAAAAGGCGCATCA 480
QY 481 GAGCTTGAAGAGAGGCGGCTTATCAAGCAAGCAAGAAATTAAGAGTGAATTAAGA 540
DB 481 GAGCTTGAAGAGAGGCGGCTTATCAAGCAAGCAAGAAATTAAGAGTGAATTAAGA 540

QY 541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGTGCTATATTAATTCGCAAG 600
DB 541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGTGCTATATTAATTCGCAAG 600
QY 601 CCTGAGTATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCTGAGTATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGCTACGGGTTGTATGACAAGACAGATGATGATGATGATGATGATGATGATGAT 720
DB 661 AGCAGCTACGGGTTGTATGACAAGACAGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACAGGTGCAATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCAATGACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTGGGACGGAAGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGGACGGAAGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATTTAGGAAAAGGCTCCATTAATTAAGAAATGAGGCTCAGCGGATCTAGTCTTC 900
DB 841 TCCGATTTAGGAAAAGGCTCCATTAATTAAGAAATGAGGCTCAGCGGATCTAGTCTTC 900
QY 901 CAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 CAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TTCAGCCAGAGATCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 TTCAGCCAGAGATCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 AATGGGCTTACACAACAATTCGAATGTGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGCTTACACAACAATTCGAATGTGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCTTTTGGTGGCGGGAATGAGACCGAAGCGGCGGACCATCATGATGACCGAGC 1140
DB 1081 ACGATCTTTTGGTGGCGGGAATGAGACCGAAGCGGCGGACCATCATGATGACCGAGC 1140
QY 1141 ACGCTTAAATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1141 ACGCTTAAATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 TATGCGGACAAATATCAACATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 TATGCGGACAAATATCAACATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 CGGATCAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 CGGATCAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GCAACCGGATTCCTCTTCTGCGGCAACATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GCAACCGGATTCCTCTTCTGCGGCAACATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TCCATGCTACACCGATGTTGCTGGAACGTGCGCACGCTTCTGTAAGCATTTTGTGAAA 1440
DB 1381 TCCATGCTACACCGATGTTGCTGGAACGTGCGCACGCTTCTGTAAGCATTTTGTGAAA 1440
QY 1441 AACGAGGCAATCAACCAAGGCTTCTATTAAGCGGCACTGATGATGATGATGATGAT 1500
DB 1441 AACGAGGCAATCAACCAAGGCTTCTATTAAGCGGCACTGATGATGATGATGATGAT 1500
QY 1501 GACATGCGCTTGTGCTACCGGAAGGATTAACAGATGATGATGATGATGATGATGAT 1560
DB 1501 GACATGCGCTTGTGCTACCGGAAGGATTAACAGATGATGATGATGATGATGATGAT 1560
QY 1561 TCCCTGAAGCTTGTGCTATGTAAGAGTCAAGTCTTCTATCAACGCAAGCAAGAGG 1620
DB 1561 TCCCTGAAGCTTGTGCTATGTAAGAGTCAAGTCTTCTATCAACGCAAGCAAGAGG 1620
QY 1621 TACTCGTTATCTGCTGCGGCAAGGCTTTGAAAATCTCCCTGGTATGATGATGCC 1680

Db	907	ATGATATACGAGTGGTGGGCTTGGAGGCTTGGCTTCCATCTTGCACAACTTATTACCCAA	966
Qy	970	GCATACAGTGTGGTGCCAGAAATTGATACAACTTCCTGGGGAGCAGCAGTAATGGGCT	1029
Db	967	GCAATTCAGTGCAGGTGGCCAGAAATTCATCTCAAACTCTGGGGGGCAGCGGTCAATGGGGCC	1026
Qy	1030	TACACAAAGCAATTCAGAAATGTGGATGAATATGTGGCCGCAAAATGATATATGATCCTT	1089
Db	1027	TACACGACGATTTCCAGAAATGTGGATGATATATGAAAGAAATGATATATGACATTCCTT	1086
Qy	1090	TTCCGTGCGGGAAATGAAAGACCGAACCGCGGAAACATCAGTGCACACGACAGCTTAA	1149
Db	1087	TTCCGGGCTGGGAAATGAAAGGCCGAAACGGCGGGTACATCAGTGACCTGTATACGGCTAA	1146
Qy	1150	AATGGAATATACAGTGGAGCTTACGAAAACTTCGCGCCAGCTTTGGGTCTTATGCGGAC	1209
Db	1147	AACGCGATACAGGTGGGCGCAACCGAAACCTGGGCTCCAGCTTCGGTTCTTATGACGAT	1206
Qy	1210	AATATCAACATGTGTGCAAGTTCCTTCACGTGGACCGACAAAGATGACGGATCAA	1269
Db	1207	AATATTAACAAGTTTGACAGTTCTCTCCGTTGGCCCGACAAAGATGGCGAATCAG	1266
Qy	1270	CCGAGTTCATGGCACCGGGAAAGTTCACTATCAGCAAGATCTTCTCTTGACCGGAT	1329
Db	1267	CCTAATGTCATGGCGGCCAGGACCAATATCTTTATCAGCAAGATCTTCTCTTGACCGGAT	1326
Qy	1330	TCCTCCTCTGGGGCAACCATGACAGTAAATATGATCATGAGGTGGAAAGTCCATGGCT	1389
Db	1327	TCCTCCTCTGGGGCAATCATGACAGCAAAATATGCTATATGGGTGGAAAGTCCATGGCA	1386
Qy	1390	ACACCGATGTTGCTGGAACGATGGCACAGCTTCGTGACATTTTGTGAAAAACAGAGC	1449
Db	1387	ACACCGATGTTGTCGGGGAAATGTTGACAGCTCGTAGACATTTTGTGAAAAATAGAGA	1446
Qy	1450	ATCACCAAGCCCTTCTCTATTAAAGCGGACATGATGGCGGTGACGTGACATGGCC	1509
Db	1447	ATCACTCTTAAGCTTCCCTATTGAAACAGCTTGAATGAGGTGTGCTGATGTTGGA	1506
Qy	1510	CTTGGCTACCCGAAACGATACCAAAGATGGGAGCGAGTGAATTTAAATCCCTGAC	1569
Db	1507	TTGGGTTATCCGAACGAAACCAAGATGGGGCGAGTACCCCTGATTAATCTTTAC	1566
Qy	1570	GTTGCTATGTGAACGAGTCCAGTTCTCTATCCACAGCCAAAGACGAGTACTGTTT	1629
Db	1567	GTTGCTATGTGAACGATCCAGTGCCCTTATCACTAGCCAAAGGACATATACCTTT	1626
Qy	1630	ACTGTACTGCGCGGAGCCCTTTGAAATCTCCCTGATATGTGATGGCCCTGCGAGC	1689
Db	1627	ACTGCAAGCGGGCGACGCCATTTAAATCTTCCCTGATATGTGATGGCCCTGCGAGC	1686
Qy	1687	ACTATGTGCTTGTATACCTCGTCAATGATTTGGATTTGGCTATTAAGCACCAACGGA	1746
Db	1750	ACACAGTATGTGGAAGAAATGACTTCTTCAGCACCATTTGACATATACATGAGTGGCCGAAT	1806
Qy	1747	ACACAGTATGTGGAAGAAATGACTTCTTCAGCACCATTTGACATATACATGAGTGGCCGAAT	1803
Db	1810	AACGTAGAAATGTATTTATTAATGACACCAAGCGGGAGGTATACAAATTGAGGTACAG	1869
Qy	1807	AACGTAGAAATGTATTTATTAATTTGGCCCAAGTGGAAATATACATTTGAGGTGCA	1866
Db	1870	GCTTATACGTAACCGGTGGACCAAGACCTTTCGTGGTGGCAATTTGAAATTA	1923
Qy	1867	GCAATATATGTGCGGTGGACCAAAACCTTCGTGGCAATTTGAACTTA	1920

RESULT 6
US-09-920-954-3
; Sequence 3, Application US/09920954
; Patent No. 6759228
; GENERAL INFORMATION:

```

? APPLICANT: TAKAIWA, MIKIO
? APPLICANT: OKUDA, MITSUYOSHI
? APPLICANT: SAKETI, KATSUHIISA
? APPLICANT: KUBOTA, HIROMI
? APPLICANT: HITOMI, JUN
? APPLICANT: KAGEYAMA, YASUSHI
? APPLICANT: SHIKATA, SHITSUJI
? APPLICANT: NOMURA, MASAFUMI
? TITLE OF INVENTION: ALKALINE PROTEASE
? FILE REFERENCE: 0327-0833-0PCT
? CURRENT APPLICATION NUMBER: US/09/920,954
? CURRENT FILING DATE: 2001-08-03
? PRIOR APPLICATION NUMBER: 09/509,814
? PRIOR FILING DATE: 2000-04-06
? PRIOR APPLICATION NUMBER: PCT/JP98/04528
? PRIOR FILING DATE: 1998-10-07
? PRIOR APPLICATION NUMBER: JP 9-274570
? PRIOR FILING DATE: 1997-06-08
? NUMBER OF SEQ ID NOS: 24
? SOFTWARE: SeqIntln version 3.0
? SEQ ID NO 3
? LENGTH: 1920
? TYPE: DNA
? ORGANISM: Bacillus sp.
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(1920)
? US-09-920-954-3

```

Query Match	76.8%	Score 1477.2;	DB 3;	Length 1920;
Best Local Similarity	85.7%	Pred. No. 0;		
Matches 1641; Conservative	0;	Mismatches 273;	Indels 0;	Gaps 0;

Qy	1	AAAAAAGSITGTTTATCTGTTTTATCAGTCGACGAGATTTGTGACCTGTGGCTA	69
Db	7	AAAGAGAGSITGTTTATCTGTTTTATCAGTCGACGAGATCTGTCACTGTGCATTA	66
Qy	70	AGTATTCACCTGCAAGTSGTSCAAGAAATTTGATCTGATTTCAAGAATTCAGCA	122
Db	67	AACATCTCTCGCTGTGTATGCAGACCTTTGATCTGATTTTAAAGAAATTCAAAC	122
Qy	130	ACAACTGATGCTAAAGTTTTCTCAAGACGGGCGAGCTGSGTGTCTGCTTTTCTG	189
Db	127	ACAACCGATGTCAGTGTGTTCTCAACACGGGCAAAAGGTGGCGCTGACATTTCTG	189
Qy	190	GAATCTGAAAATGTGAACTCCCAAAAGSITTTGCAAGAAAGCTTTGAAACGTCCG	249
Db	187	GAATCTGAAAATGTGAACTCTTAAAGATTTGCTAAAGAACTTGAAACGTACCG	246
Qy	250	AATATATTAATCTCATATTAATCCATTAATCAATGACCAATTTTAAAGAAACAAAC	309
Db	247	AATATATTAATCTCATATTTGTCCATTTCAATGGCCCATTTTAAAGAAACAAAC	306
Qy	310	CTGGAAGAAACAGGGGCAAGATTTCTGACTACATCTGTATTTGCTTACATTTG	363
Db	307	CTGAGACAACTGAGCAAAAGATTTCTGACTACATCTGTATTTGCTTATTTG	366
Qy	370	TATGAGGGCGATTTAATCAGCAACAGCACATTTGACCGTGAATCCGTGAGCT	422
Db	367	TATGAGGGGGAATTTCAATCAAAAGTCCGCTCCATTTGAAACCTGTGAATCA	426
Qy	430	TATTTGCGATTTATCAGATATGATCTCCAGCTTTTCAAAAGGGGCAATCAGAC	489
Db	427	TACTTCCGAAATTCAAATATGATCTCCAGCTTTTCAAAAGGGGCAATCAGAC	486
Qy	490	AAAGCAGTGGCTTGATACAAAGCAAGAAATTAAGAGTGCATTTAAGAGCATGAA	549
Db	487	AAAGGTTGGCTTGATACAGAGCAAGAAATTAAGAGTGCATTTAAGAGCATGAG	546
Qy	550	CAATGTGCATTTGCCATAAGCAATATGTCTATATATACGCAAAAGCTGAGTAT	609
Db	547	GAATGTGCATATGATGTGCAAGCAATATGCTCATATATATACGCAAAAGCTGAGTAT	606

QY	610	AAGGGAATGAATGAGTTGCGGTGGAAATGGTCAAGCGGAGTGGCTCAGAGCATAC	666
Db	607	AAGGTGATTAATGAATGTGGCCAGAGGTATTGTCAAGCGGAGTGTGGCAAGACAGCTAC	666
QY	670	GGGTGTATGAGCAAGAGGACAGATCTGATGGCGTTTGCATACAGGCGTTGATACAGTCCG	729
Db	667	GGTTTGTATAGCAAGAGGCCAGATTTGTTCGAGTTTGCAGATTGTGGATTGATACAGGAATA	726
QY	730	AATGACAGTTCCATGTCATGAAGCTTTCCGCGGGGAAAATTACTGCAATTATATGCAATTGGGA	789
Db	727	AACGACAGTTTCATGTCATGAAGCTTTCCCGGTAATAATTAACAGCACTAATATGCACTGGGT	786
QY	790	CGGAGGAAATAAATGGCCAAATGATACGAATGGTCATATGTAACCAATGGTCGGCTCCGATTA	849
Db	787	CGGACGAATAAATGGCCGAATATACGAACGGTCATGTATCCCATATGGCAAGTTCGGATTA	846
QY	850	GGAATACGGCTCCACTAATATAAGGAATGGCGCTCAGAGCGAATCTAATCTTCCATCTATC	909
Db	847	GGAATATGGGCAACGAATTAAGGAATGGCACTCAGAGCGAATCTGTTTTTCAATCATTAC	906
QY	910	ATGATATGGCGTGGGGGAACTTGGAGGACTAATCTTGCAATCTGCAAACTTTATTCAGCCAA	969
Db	907	ATGATATGACGATGTGTGGCTTGGAGGCTTGCTTCAATCTGCAAACTTTATTCAGCCAA	966
QY	970	GCATACAGTGTGTGTGCCGAATTTCAATACAAATCTCTGGGGAGCAGCACTGAATGGGGCT	1022
Db	967	GCATTTACAGTCAGGTGTCCAGAAATTCATACAAATCTCTGGGGGAGCAGCGTGAATGGGGGC	1022
QY	1030	TACACAAACAGATTTCCAGAAATGTGATGTACTATGTGCGGAAAATGATATGACGATCTT	1088
Db	1027	TACAGCAACGATTTCCAGAAATGTGATGTACTATGTAAAGAAAATGATATGACGATCTT	1088
QY	1090	TTTCGCTGTCCGGGAATGAAGAACCGAAACCGGCGGAACCATGATGACACAGGCAACGCTAAA	1144
Db	1087	TTTCGCGGCTGGGAATGAAGGCCGAACCGGCGGATACATGATGACACCTGTGATACGCTAAA	1144
QY	1150	AATGCAATTAACGTGCGAGCTACGGAATACTCCGCCAAGCTTTGGGTCTTATGGGAC	1200
Db	1147	AACGCAATTAACGTGCGGCGCAACCGGAATACTGCGCTCAAGCTTCGATTCATATGCAAGAT	1200
QY	1210	AATATCAACCATGTGGCAAGTCTCTTACGATGAGCGGCAAAAGGATGGAAGGATCAAA	1266
Db	1207	AATATTAACCATGTTGCAAGTCTCTTCCCGTGGCCGACAAAAGATGGGCGAATCAAG	1266
QY	1270	CCGATGTCTATGGCAACCGGGAACGTTCAATCACTATCAGCAAGATCTTCTTTGACCCGAT	1332
Db	1267	CCGATGTCTATGGCGGCAAGGACATCACTTTATCAGCAAGATCTTCTTTGACCCGAT	1332
QY	1330	TTCCTCTTTTGGGCGCAACATGACAGTAATATATGATACATGGGTGGAACGTTCATGGCT	1388
Db	1327	TTCCTCTTTTGGGCGGATCATGACAGCAATATGTCTTAATATGGTGTGGAACGTTCATGGCA	1388
QY	1390	AACCGGATGTGTGCTGGAAACGATGCGGACAGCTTCGTAACATTTTGTGAAAACAGAGGC	1444
Db	1387	AACCGGATGTGTGCGGGGAATGTGACAGCTTCGTAACATTTTGTGAAAATAGAGGA	1444
QY	1450	ATCACAACCAAGCTTCTCTATTAATAAGCGGACATGATTTGCGGTGCACTGACATCGGC	1500
Db	1447	ATCAGCTCTTAAGCTTCTCTATTAAGAAGACGTTGATTTGCAAGGTCTCTGATGTGGA	1500
QY	1510	CTTGGCTACCCGAACGGTAAACCAAGATGGGGAAGAGTGAATTTGGATTAATCTCTGAAC	1566
Db	1507	TTGGGTTATCCGAACGGAAACCAAGATGGGGCGAGTGAACCTGTGATTAATCGTTGAAC	1566
QY	1570	GTTCGCTATGTGAACGATTCAGTTCTCTATCCACCAAGCCAAAAGCGAATGATCTGTTT	1622
Db	1567	GTTCGCTATGTGAACGAATTCAGTGTCTCTATCAATGACCAAAAAGCGACATATACCTTT	1622
QY	1630	ACTGCTATCGCGGCAAGCTTTGAAAATCTCCCTGTATGTGCTGATGCCCTCGAGAC	1688
Db	1627	ACTGCAACGGCGGCAAGCATTTGAAAATCTCCCTGTATGTGCTGATGCCCTCGAGAC	1688
QY	1690	ACAACTGCTTCGTAACGCTTGTCAATGATCTGGAACCTTGTCATTAACGCTCCAAATGGC	1744

QY	DB	1687	ACTACTGCTTCTGTAACCTTGGTCAATATTTGGATTGGTTCATTACAGCACCAACGGA	1744
QY	1750	ACACAGATGTAGGAATGACTTTACTTCCCATACAAATGATTACTGGAGTGGCCGAAT		1805
DB	1747	ACAAGATATGTGCGGAATGACTTCTCAGCACCACTTTGACAAATTAACCTGGGATGGCCGAAT		1806
QY	1810	AACGTAGAAATGATTTTATTATATGACACCAAAAGCGGAGGTATACAAATTGAGGTACAG		1866
DB	1807	AACGTAGAAATGATTTTATTATATGACACCAAAAGGTATACAAATTGAGGTACAG		1866
QY	1870	GCTTATACGTACCGGTGGACCAAGACACTTCTCGTTGGCAATTTGAAATTTAA	1923	
DB	1867	GCATTTATGTGCGCGTGGACCAACAAACTTCTCGTTGGCAATTTGAACTTAA	1920	
RESULT 7				
US-08-873-479-41				
; Sequence 41, Application US/08873479				
; Patent No. 5891701				
; GENERAL INFORMATION:				
; APPLICANT: Sionia, Alan				
; APPLICANT: Lyne, Christlanson				
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide				
; TITLE OF INVENTION: Haying Protease Activity				
; NUMBER OF SEQUENCES: 57				
; CORRESPONDENCE ADDRESS:				
; ADDRESSER: No. 5891701o No. 5891701disk of No. 5891701th America				
; STREET: 405 Lexington Avenue				
; CITY: New York				
; STATE: NY				
; COUNTRY: USA				
; ZIP: 10174				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Diskette				
; COMPUTER: IBM Compatible				
; OPERATING SYSTEM: DOS				
; SOFTWARE: FaecSeq for Windows Version 2.0				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/873.479				
; FILING DATE: 12-JUN-1997				
; CLASSIFICATION: 530				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Agt18, Cheryl H				
; REGISTRATION NUMBER: 34,086				
; REFERENCE/DOCKET NUMBER: 5251.000-US				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 212-867-0123				
; TELEFAX: 212-878-9655				
; TELEX:				
; INFORMATION FOR SEQ ID NO: 41:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 3003 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: Genomic DNA				
US-08-873-479-41				
Query Match				
Best Local Similarity 60.7%; Score 1166.4; DB 2; Length 3003;				
Matches 1453; Conservative 0; Mismatches 456; Indels 3; Gaps 1				
QY	12	GAAGAAAGGTGTTTATCTGTTTATCAGTCGACAGCAATTTGTGACCTGGTGGCTTAAG	71	
DB	863	GAAGAGGATTTTATTCGTTTATATCGTTTATATCGTTGCTGACATATTTGTTCTGTGGCTTTAAG	922	
QY	72	TATTCATCTGCAAGGTGTGCAAGCAATTTGATCTGATCTGATTTCAAGCAATTCAGACAC	131	
DB	923	CAGTCTTCTTACTATGTGGGCGAACAATTTGAAATTTGACCTTTAAGGGGATGAGACACT	982	
QY	132	AACGTATCTAAAGTTTCTCGCAAGGAGGGGCAAGCTGTGTGCTGCTTTTCTGTGGA	191	

Db 983 TACGTAGAGAGGCTGCCAACAAGAGAAAAAGCGAAGGCAATCTTTCTTGTA 1042
Qy 192 ATCGAATATGTAAATCTCCAAAAGTTTGAGAGAGCTTGAAACATCTCCGGAAA 251
Db 1043 CTCTGAAATGTGAAATCCCAAGAGATTTCAAAACAACTAGAGATGTTCCAGGGA 1102
Qy 252 TAAATATCTCATATATTCATATTCATATTCATATTCATATTCATATTCATATTC 311
Db 1103 TAAAGCTATATATGCTTCAATTTGACGACCTATTTAGAGAAACGCACTTCACT 1162
Qy 312 GAAAAAACAAGGCGCAAGATTTCTGCACTATCTGATTTATGCTTATCTGAGTA 371
Db 1163 AGGAAGACCGGAGCGAATAATCTGATTCATACAGATTACGCTTATATTCGATA 1222
Qy 372 TGAGGGGATATATGCTACCAACAAGCACAATTGAGCAGTGAATCCGTGAGCTTA 431
Db 1223 TGATGGGATGTAAAGGCGGTAACTAACGAAATTCGCAATTTGGAATCGGTGAACA 1282
Qy 432 TTTGCGATATATACAGATATGATCCCGAGCTTTTCAAAAAGGCGCATCAGAGCTTGA 491
Db 1283 TTTACCTTATATATATATAGACCGCAATTTATTTCCAGAGAGCTTTCTGAATTTAGTA 1342
Qy 492 AGCAGTGGCGCTTGATACAAAGCAAAAATAAGAGTGCATTTAAGAGCATCGACA 551
Db 1343 AACAGTAGCTTTAGATTAAGAAAGCAAGAGTAAGAGTACGTTTAAGAGATTGGAACA 1402
Qy 552 AATCGCAATTTGGCAATAGCAATATGCTATATATTTACGCAAGAGCTTGAATTA 611
Db 1403 AATTTGCCAATAGCGCAAAATATATGATTTATATGTAACCCCAAGCTCAATATGCA 1462
Qy 612 GGTATGATATGATGTTGCGGTAATTTGCAAAAGCGATGTTGCTGAGAGCTACGG 671
Db 1463 AGTTTGAATGAGCTGCGCCGTGATTTGAAAGCAGAGCTGCAAAAATTAATTTGG 1522
Qy 672 GTTGTATGACAAAGACAGATCGTACGATTTGCGATACAGGCTTGATACAGTGCACA 731
Db 1523 CTATATGACAAAGACAGATTTGATGAGATTTGCTGATACGAGCTTGAATACAGAGAAA 1582
Qy 732 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
Db 1583 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1642
Qy 792 GACGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
Db 1643 AACGATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1702
Qy 852 AAACGGCTCACTAATTAAGAAATGCGCTCAGGCAATCTAGCTTTCATCTATCAT 911
Db 1703 AATATG---CTACAAATTAAGGATGAGACCGCAAGCAATCTAGCTTTCATCTATCAT 1759
Qy 912 GGTATGCGGTGGGGAATTTGAGAGATCTACCTTCAATCTGCAATCTTATTCAGCAAGC 971
Db 1760 GGTATGCGGTGGGGAATTTGAGAGATCTACCTTCAATCTTATTCAGCAAGC 1819
Qy 972 ATACGATGCGGTGGGGAATTTGAGAGATCTACCTTCAATCTTATTCAGCAAGC 1031
Db 1820 ATATATGCTGAGGAGCAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1879
Qy 1032 CACAACAGATTTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
Db 1880 TACGACAGATCTCCGAAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1939
Qy 1092 CGCTGCGGGAATGAGAGCGAAGCGGGAACCATCATGTCACAGGCAAGCTTAATAA 1151
Db 1940 TCGGCGCGGGAATGAGAGCGAAGCGGGAACCATCATGTCACAGGCAAGCTTAATAA 1999
Qy 1152 TCGAATTAACAGTGGAGCTACGGAATACTCCGCGCAAGCTTGGGCTTATACCGACAA 1211
Db 2000 TCGAATTAACAGTGGAGCTACGGAATACTCCGCGCAAGCTTGGGCTTATACCGACAA 2059
Qy 1212 TATCAACAGTGGAGCTTCTTCAAGTGGAGCGCAAAAGATGAGCGATCAAC 1271
Db 2060 TATTAACAGTGGAGCTTCTTCAAGTGGAGCGCTTCTTCAAGATGAGCGATTAAGCC 2119

Qy 1272 GATGTCATGCGACCGGGAAGCTTCACTATGAGAGATCTTCTTTCACCGGATTC 1331
Db 2120 GAGGTCATGCGACCGGGAAGCTTCACTATGAGAGATCTTCTTTCACCGGATTC 2179
Qy 1332 CTCCTTCTGGGCAACATGACAGTAAATATGATATCATGAGTGAACGTCATGAGCTAC 1391
Db 2180 CTCATTTCTGGGCAACATGATATGATATGATATGATATGATATGATATGATATGATAT 2239
Qy 1392 ACCGATGCTTCTGGAACGTCGACAGCTTCTGAGCATTTTGTGAAAAACAGAGCAT 1451
Db 2240 TCCAAATTTGACAGTAAATGTTGCACAATTAAGGAGCATTTTGTGAAAAATGAGGGGT 2299
Qy 1452 CACACCAAGCCCTTCTCTATTAACACCGGCACTGATTTCCGCTGACGTCATTCGGCT 1511
Db 2300 AACTCTTAAGCTTCTCTTAAACCTGCTTAAATTTGACAGGCTGCGGATGTTGACT 2359
Qy 1512 TGGCTACCGGAAGCTTAACCAAGATGGGACAGTGAATGATTAATCCCTGAACGT 1571
Db 2360 TGGCTTTCCAATGATTAACCAAGATGGGAGAGTAACTTGAATTAATCCCTGAATGT 2419
Qy 1572 TGGCTATGGAACGATTCAGTTCCTATTCACAGCCCAAAAAGCAAGTACTGTTTAC 1631
Db 2420 CGCATTTGGAATGAACAGAGCCCTTATCAACAGTCAAAAGCAACATATTCGTTTAC 2479
Qy 1632 TGTCTATGCGCGGACGCTTTGAAAATCTCCCTGATATGATGATGATGATGATGATGAT 1691
Db 2480 GGTCTAAGCTGATTAACCTCTTAAATATCACTGTTTGTGATGATGATGATGATGATGAT 2539
Qy 1692 AACTGCTCGTAAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1751
Db 2540 GACGATCACTAATCTTGTGATATGATTAATGATTAATGATTAATGATTAATGATTAAT 2599
Qy 1752 ACAGATGTAAGAAATGATTAATCTTGCATTAACATGATTAATGATTAATGATTAATGAT 1811
Db 2600 TAAATAGCTCGGAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 2659
Qy 1812 CTTAAGAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1871
Db 2660 CTTAAGAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 2719
Qy 1872 TTTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1923
Db 2720 TTTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 2771

RESULT 8
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHREIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:


```
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A., 768
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-9300
/ TELEFAX: (703) 683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: PTZ9pt-P18
/ US-08-232-463-14

Query Match 2.8%; Score 54; DB 2; Length 7218;
Best Local Similarity 2.4%; Pred. No. 8.5e-06;
Matches 9; Conservative 222; Mismatches 147; Indels 0; Gaps 0;

QY 468 AAAAGGCGCATGACGCTGTAAAGCAAGTGGCGCTTGATACAAAGCAAAAAATTAAGA 527
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1430 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1371

QY 528 GGTGCAATTAAAGGCGATGGAACAATTCGACCAATTCGCAATTAAGCAATGATGCTGATA 587
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1370 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1311

QY 588 TATTACGGAAGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1310 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1251

QY 648 GGATGTGGCTCAGACGAGTACGCGTTGATGATGATGATGATGATGATGATGATGATG 707
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1250 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1191

QY 708 TACAGGCGTTGATACAGCTGCAATGACAGTGTGATGATGATGATGATGATGATGATG 767
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1190 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1131

QY 768 TACTGATTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 827
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1130 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1071

QY 828 GCATGTGGCTGCTCGCT 845
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1070 RRRRATCGMAAGCTCCT 1053

RESULT 9
US-09-621-976-2574
/ Sequence 2574, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Mline Edwards, J. B.
/ APPLICANT: Jobert, S. Y.
/ APPLICANT: Giordano, J. Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 2574
/ LENGTH: 834
/ TYPE: DNA
/ ORGANISM: Homo sapiens
```

```
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 497..826
/ US-09-621-976-2574

Query Match 2.2%; Score 42.4; DB 3; Length 834;
Best Local Similarity 16.9%; Pred. No. 0.012;
Matches 28; Conservative 81; Mismatches 57; Indels 0; Gaps 0;

QY 822 TGTACGCAATGTGCGTCTCGTATTAAGAAACGCGCTCCACTAATTAAGAAATGGCGCC 881
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 38 WSGCRSTSRMRSSKSYGRACMSRRGRKACMSWTGMSCTGRRMSWKRKRSSWYSTK 97

QY 882 TCAGCGCAATCTAGCTTCCTCCAACTATCATCATGATGATGATGATGATGATGATGATG 941
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 98 WKMRGYSWTYSYRYMRCSYMMWMMCMSCYKRSRMCMMGSGRSMCMYIGMSAATCRVC 157

QY 942 TTCGAATCTGCAACCTATTACAGCAAGCATTAAGTGTGCTGTC 987
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 MRYRMTCAKSRKMKMKMKMSGMSCKMSCKMSCKMSCKMSCKMSCKMSCKMSCKMSCK 203

RESULT 10
US-09-620-312D-590/C
/ Sequence 590, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyun
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yungling
/ APPLICANT: Wang, Duntui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ FILE REFERENCE: Polypeptides
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 590
/ LENGTH: 5053
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (154) ..(2409)
/ US-09-620-312D-590

Query Match 2.2%; Score 41.8; DB 3; Length 5053;
Best Local Similarity 57.1%; Pred. No. 0.056;
Matches 76; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 174 TGTGCTTTTGTGTGATCTGAATAATGTGAATCCCAAAAGTTTGCAGAAAGAGCT 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 3158 TGGAGTTATGTGTGATGAACAAGCTACTGAACACCTGCAATGCTGTTTAAAGTTT 3099

QY 234 TGAACAGTCCCGGCAAAATTAATTAATCTCATATTAATCAATTCAGCAATTTTGA 293
```


Db 391 VAMMYSRDTNTDMMWMTSDMBWHYTVDTYTMGRANNNNNNNRBCKTSSMMWMDHM 332
 Qy 362 TTGTGAGATATAGGCGCATGTTAAGTCGCAACGACCATGAGCATCG 421
 Db 331 NTHCTYGNNTWGSABMAMSMWAGASBYTYNNCMRTYNGKTMNNNNNK-AMYR 273
 Qy 422 TGGACCTTATTTGCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGCGATCG 481
 Db 272 TKTIVAMCNRNRYTDTAVTWBKNNYKCYAYBYTYRMYKHHMBWRBAPHSSNNMYK 213
 Qy 482 AGCTTGTAAGACGAGTGGCGCTTGATACAAACGAGAAAATAAGAGTGCAATTAAAG 541
 Db 212 CNKTYVSVNHYAHMYBKWABAIVGCNNNNKDRMHHHMCATNNNNMMWYAYVHHHKKG 153
 Qy 542 GCGTGAACAATCCGACAAATTCGCAATAGCAATGATGCTATATATTACGCGAAAC 601
 Db 152 KRAMNNNTKABRDDHBAHVKTYWRYDYWCAMCMMNAKAVRTAKHMYTTRVVSAN 93
 Qy 602 CTGAGTATAGGTGATGATGAT 624
 Db 92 NTGVAMMMRMKCMWYSMMNNRY 70

RESULT 14
 US-08-750-532-2
 ; Sequence 2, Application US/08750532
 ; Patent No. 5756339
 ; GENERAL INFORMATION:
 ; APPLICANT: MITTA, Masanori
 ; APPLICANT: YAMAMOTO, Katsuhiko
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: TSUNASAWA, Susumu
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENE
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
 ; STREET: 419 Seventh Street N.W., Suite 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: United States of America
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/750,532
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP95/01095
 ; FILING DATE: 05-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 1994/130236
 ; FILING DATE: 13-JUN-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 1994/173912
 ; FILING DATE: 26-JUL-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BROWDY, Roger L.
 ; REGISTRATION NUMBER: 25,618
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-5197
 ; TELEFAX: (202) 737-3528
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2835 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single

; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-750-532-2
 Query Match 2.0%; Score 37.6; DB 2; Length 2835;
 Best Local Similarity 54.3%; Pred. No. 0.86;
 Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 1037 CAGATCCGAAGAAATGATGATGATGCGCCGAAAATGATATGACATCTTTTCGCTG 1096
 Db 1529 CAGAAAGGTTGCTGTGATGATGATGCTTACGAAAAGTACGGTGTGATTCGTAATAGCTG 1588
 Qy 1097 CCGGAATGAAGAACCGAACGCGGACCATCAGTGCACGACGACAGCTAAATGCA 1156
 Db 1589 CAGGAATGAAGGTCCTGCGCATTAACATGTTGGAAGTCTGTGTTGCAACAAAGGCA 1648
 Qy 1157 TAAAGTCGAGCTACGGA 1176
 Db 1649 TAACTGTGAGCGCTGCA 1668

RESULT 15
 US-08-750-532-8
 ; Sequence 8, Application US/08750532
 ; Patent No. 5756339
 ; GENERAL INFORMATION:
 ; APPLICANT: MITTA, Masanori
 ; APPLICANT: YAMAMOTO, Katsuhiko
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: TSUNASAWA, Susumu
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENE
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
 ; STREET: 419 Seventh Street N.W., Suite 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: United States of America
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/750,532
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP95/01095
 ; FILING DATE: 05-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 1994/130236
 ; FILING DATE: 13-JUN-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 1994/173912
 ; FILING DATE: 26-JUL-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BROWDY, Roger L.
 ; REGISTRATION NUMBER: 25,618
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-5197
 ; TELEFAX: (202) 737-3528
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4765 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: April 8, 2006, 01:34:45 ; Search time 1380.58 Seconds
(without alignments)
11518.383 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923

Sequence: 1 atgagagaagaagaagaagt.....cgttgcaatcgtgaattaa 1923

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:*

1: /cgn2_6/prodata/1/pubpna/us07_PUBCOMB.seq:*

2: /cgn2_6/prodata/1/pubpna/us08_PUBCOMB.seq:*

3: /cgn2_6/prodata/1/pubpna/us09_PUBCOMB.seq:*

4: /cgn2_6/prodata/1/pubpna/us10_PUBCOMB.seq:*

5: /cgn2_6/prodata/1/pubpna/us10A_PUBCOMB.seq:*

6: /cgn2_6/prodata/1/pubpna/us10C_PUBCOMB.seq:*

7: /cgn2_6/prodata/1/pubpna/us10D_PUBCOMB.seq:*

8: /cgn2_6/prodata/1/pubpna/us10E_PUBCOMB.seq:*

9: /cgn2_6/prodata/1/pubpna/us10F_PUBCOMB.seq:*

10: /cgn2_6/prodata/1/pubpna/us11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1923	100.0	1923	3	US-09-920-954-5
2	1923	100.0	1923	7	US-10-456-479-3
3	1923	100.0	1923	7	US-10-784-870-5
4	1923	100.0	1923	8	US-10-820-712A-2
5	1923	100.0	1923	9	US-10-820-714A-2
6	1923	99.8	1923	3	US-09-920-954-7
7	1919.8	99.8	1923	7	US-10-784-870-7
8	1477.2	76.8	1920	3	US-09-920-954-3
9	1477.2	76.8	1920	7	US-10-784-870-3
10	1305	67.9	1305	6	US-10-385-662-1
11	1281.8	66.7	1302	8	US-10-820-712A-13
12	1281.8	66.7	1302	9	US-10-820-714A-14
13	997	51.8	1302	8	US-10-820-712A-11
14	997	51.8	1302	9	US-10-820-714A-12
15	854.6	44.4	1302	8	US-10-820-712A-21
16	854.6	44.4	1302	9	US-10-820-714A-22
17	740.4	38.5	1299	8	US-10-820-712A-19
18	740.4	38.5	1299	9	US-10-820-714A-20
19	738.8	38.4	1299	8	US-10-820-712A-15
20	738.8	38.4	1299	9	US-10-820-714A-16
21	737.2	38.3	1299	8	US-10-820-712A-17
22	737.2	38.3	1299	9	US-10-820-714A-18
23	48	2.5	54	6	US-10-385-662-20

C	24	45	2.3	415	5	US-10-016-349A-101	Sequence 101, App
C	25	41.8	2.2	1095	5	US-10-106-698-1076	Sequence 1076, App
C	26	41.8	2.2	1297	3	US-09-822-846-311	Sequence 311, App
C	27	41.8	2.2	4440	9	US-10-756-149-4271	Sequence 4271, App
C	28	41.8	2.2	5053	5	US-10-037-270-590	Sequence 590, App
C	29	41.8	2.2	5053	6	US-10-117-722-590	Sequence 590, App
C	30	41.8	2.2	5053	9	US-10-122-851-590	Sequence 590, App
C	31	41.8	2.2	8195	3	US-09-764-891-9013	Sequence 9013, App
C	32	40.2	2.1	604	4	US-09-925-065A-3789	Sequence 3789, App
C	33	40	2.1	834	5	US-10-123-155-148	Sequence 148, App
C	34	40	2.1	834	6	US-10-146-731-148	Sequence 148, App
C	35	40	2.1	834	6	US-10-140-672-148	Sequence 148, App
C	36	40	2.1	834	6	US-10-141-761-148	Sequence 148, App
C	37	40	2.1	834	6	US-10-142-885-148	Sequence 148, App
C	38	40	2.1	834	6	US-10-158-790-148	Sequence 148, App
C	39	40	2.1	834	6	US-10-137-871-148	Sequence 148, App
C	40	40	2.1	834	6	US-10-140-923-148	Sequence 148, App
C	41	40	2.1	834	6	US-10-141-756-148	Sequence 148, App
C	42	40	2.1	834	6	US-10-141-759-148	Sequence 148, App
C	43	40	2.1	834	6	US-10-140-805-148	Sequence 148, App
C	44	40	2.1	834	6	US-10-140-864-148	Sequence 148, App
C	45	40	2.1	1330	9	US-10-956-157-281	Sequence 281, App

ALIGNMENTS

RESULT 1									
US-09-920-954-5									
Sequence 5, Application US/09920954									
Publication No. US20020064854A1									
GENERAL INFORMATION:									
APPLICANT: TAKAIWA, MIKIO									
APPLICANT: OKUDA, MITSUYOSHI									
APPLICANT: SAKKI, KATSUSHI									
APPLICANT: KUBOTA, HIROMI									
APPLICANT: HITOMI, JUN									
APPLICANT: KAGEYAMA, YASUSHI									
APPLICANT: SHIKATA, SHITSUMI									
APPLICANT: NOMURA, MASAFUMI									
TITLE OF INVENTION: ALKALINE PROTEASE									
FILE REFERENCE: 0327-0832-0PCT									
CURRENT APPLICATION NUMBER: US/09/920, 954									
CURRENT FILING DATE: 2001-08-03									
PRIOR APPLICATION NUMBER: 09/509, 814									
PRIOR FILING DATE: 2000-04-06									
PRIOR APPLICATION NUMBER: PCT/JP98/04526									
PRIOR FILING DATE: 1998-10-07									
PRIOR APPLICATION NUMBER: JP 9-274570									
PRIOR FILING DATE: 1997-06-08									
NUMBER OF SEQ ID NOS: 24									
SOFTWARE: PatentIn version 3.0									
SEQ ID NO 5									
LENGTH: 1923									
TYPE: DNA									
ORGANISM: Bacillus sp.									
FEATURE:									
NAME/KEY: CDS									
LOCATION: (1)..(1923)									
US-09-920-954-5									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGAGAAAGAAAGGATTTTATCTGTTTATCGCTGCACGATTTGTGACT	60						
Db	1	ATGAGAAAGAAAGGATTTTATCTGTTTATCGCTGCACGATTTGTGACT	60						
Qy	61	GTTGCGTTAAGTATCATCTGCAGGTGTCAGGAATTTGATTCGATTTCAAGA	120						
Db	61	GTTGCGTTAAGTATCATCTGCAGGTGTCAGGAATTTGATTCGATTTCAAGA	120						

SEQ ID NO 3
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
OTHER INFORMATION:
US-10-456-479-3

Query Match 100.0%; Score 1923; DB 7; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATCAAGTGCAGGATTTTGTCACT 60
DB 1 ATGAGAAAGAAAGAGTGTGTTTATCTGTTTATCAAGTGCAGGATTTTGTCACT 60
QY 61 GTTGCCTTAAGTAATCCATCTGCAGGTGTCAGAGAAATTTGATCTGAATTTCAAGGA 120
DB 61 GTTGCCTTAAGTAATCCATCTGCAGGTGTCAGAGAAATTTGATCTGAATTTCAAGGA 120
QY 121 ATTCAAGCAACAACTGATGCTAAAGTTTCTCCAGAGGCGCAGACTGTGCTGCT 180
DB 121 ATTCAAGCAACAACTGATGCTAAAGTTTCTCCAGAGGCGCAGACTGTGCTGCT 180
QY 181 TTTCTGTGGAATCTGAAAATGTGAACTCCCAAAAGTTTGCAGAGAAAGCTTGAAACA 240
DB 181 TTTCTGTGGAATCTGAAAATGTGAACTCCCAAAAGTTTGCAGAGAAAGCTTGAAACA 240
QY 241 GTCCCGCAATATAAATCCAAATATCCAAATTCGAATGGAACCAATTTTGAAGAAACA 300
DB 241 GTCCCGCAATATAAATCCAAATATCCAAATTCGAATGGAACCAATTTTGAAGAAACA 300
QY 301 AAACAGAGCTGGAAGAAAGAGGCGCAAGATTTCTGCACTACATACCTGATTAAGCTTAC 360
DB 301 AAACAGAGCTGGAAGAAAGAGGCGCAAGATTTCTGCACTACATACCTGATTAAGCTTAC 360
QY 361 ATTGTGAGATTAAGGCGCAATTTAAGTCAGCAACAGCAACATTGACACGTGGAATCC 420
DB 361 ATTGTGAGATTAAGGCGCAATTTAAGTCAGCAACAGCAACATTGACACGTGGAATCC 420
QY 421 GTGAGGCTTAATTTGCGGATATACAGAAATGATCCCGAGCTTTTCAAAAAGGCGCAACA 480
DB 421 GTGAGGCTTAATTTGCGGATATACAGAAATGATCCCGAGCTTTTCAAAAAGGCGCAACA 480
QY 481 GAGCTTGTAAAGCAGTGGCGCTTGAATCAAGCAAGAAATTAAGAGTGCATTTAAGA 540
DB 481 GAGCTTGTAAAGCAGTGGCGCTTGAATCAAGCAAGAAATTAAGAGTGCATTTAAGA 540
QY 541 GGCATGGAACAAATCCGCAATTCGCAATAGCAATGATGTCTATTAATTAACGGCAAG 600
DB 541 GGCATGGAACAAATCCGCAATTCGCAATAGCAATGATGTCTATTAATTAACGGCAAG 600
QY 601 CCTGATATTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCTGATATTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGCTACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 AGCAGCTACGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACAGGTGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGGTGCAATGACAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCAATGGAACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCAATGGAACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGTATTAGAAACGCTCTCACTAATTAAGAAATGGCGCTCAGCGGAAATCTAGCTTC 900
DB 841 TCCGTATTAGAAACGCTCTCACTAATTAAGAAATGGCGCTCAGCGGAAATCTAGCTTC 900

QY 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGAGAGACTACCTTGAATCTGCAAACTTTA 960
DB 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGAGAGACTACCTTGAATCTGCAAACTTTA 960
QY 961 TTGAGCGCAAGCATACAGTGTGGGTCAGAAATTTCAAACTCTGAGGAGCGAGCTG 1020
DB 961 TTGAGCGCAAGCATACAGTGTGGGTCAGAAATTTCAAACTCTGAGGAGCGAGCTG 1020
QY 1021 AATGGGCTTACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 AATGGGCTTACACACAGATTCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ACGATCTTTTCCGTGCGGGAATGAAGAACCGAACCGGCAACCATCAGTGCACAGGC 1140
DB 1081 ACGATCTTTTCCGTGCGGGAATGAAGAACCGAACCGGCAACCATCAGTGCACAGGC 1140
QY 1141 ACAGCTAAATGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 ACAGCTAAATGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 TATGCGCAATATCAACATGTCAGACTTCTTCAAGTGCACGCAAGAGATGA 1260
DB 1201 TATGCGCAATATCAACATGTCAGACTTCTTCAAGTGCACGCAAGAGATGA 1260
QY 1261 CGGATCAACCGGATGTCATGCAACCGGAAAGCTTCACTATCAAGAAATCTTCTT 1320
DB 1261 CGGATCAACCGGATGTCATGCAACCGGAAAGCTTCACTATCAAGAAATCTTCTT 1320
QY 1321 GCAACCGGATCTCTCTGTCGCGGAAACGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GCAACCGGATCTCTCTGTCGCGGAAACGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TCCATGCTTACACCGATTCGTCGGAAGAGTGCACAGCTTGTGAGCAATTTGTGAAA 1440
DB 1381 TCCATGCTTACACCGATTCGTCGGAAGAGTGCACAGCTTGTGAGCAATTTGTGAAA 1440
QY 1441 AACAGGCAATCAACCAAGCTTCTTATTAAGAGCGCACTGATTCGCGTGACGT 1500
DB 1441 AACAGGCAATCAACCAAGCTTCTTATTAAGAGCGCACTGATTCGCGTGACGT 1500
QY 1501 GACATGCGCTTGGCTAACCGGAACGGTAACCAAGATGAGGAGAGATGATGATGATGATGAT 1560
DB 1501 GACATGCGCTTGGCTAACCGGAACGGTAACCAAGATGAGGAGAGATGATGATGATGATGAT 1560
QY 1561 TCCCTGAACGTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1561 TCCCTGAACGTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 TACTGCTTATCTGCTACGCGGCAAGCTTTGAAAATCTCCCTGATATGATGATGATGATG 1680
DB 1621 TACTGCTTATCTGCTACGCGGCAAGCTTTGAAAATCTCCCTGATATGATGATGATGATGAT 1680
QY 1681 CCTGCGAGCAACATCTCTTCCGTAACGCTTGTCAAGATGATGATGATGATGATGATGAT 1740
DB 1681 CCTGCGAGCAACATCTCTTCCGTAACGCTTGTCAAGATGATGATGATGATGATGATGAT 1740
QY 1741 CCAATGCGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 CCAATGCGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GGCAGCAATTAAGTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1801 GGCAGCAATTAAGTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 GAGGTACAGCTTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1920
DB 1861 GAGGTACAGCTTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 3
US-10-784-870-5
Sequence 5, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: OKUDA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKET, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
PRIORITY FILING DATE: 2004-02-24
PRIORITY APPLICATION NUMBER: US/09/509,814A
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: PCT/JP98/04528
PRIORITY FILING DATE: 1998-10-07
PRIORITY APPLICATION NUMBER: JP 9-274570
PRIORITY FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1923)
US-10-784-870-5

Query Match 100.0%; Score 1923; DB 7; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCGCTGACGCAATTTTGTGACT 60
1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCGCTGACGCAATTTTGTGACT 60
61 GTTGCTTAAGTATCATCTGAGGTGTCAGAGAAATTTGATGTGAATTTCAAAGA 120
61 GTTGCTTAAGTATCATCTGAGGTGTCAGAGAAATTTGATGTGAATTTCAAAGA 120
121 ATTGAGACAACTGATGCTAAAGGTTTCTCCAGCAGGCGACAGCTGTGCTGCT 180
121 ATTGAGACAACTGATGCTAAAGGTTTCTCCAGCAGGCGACAGCTGTGCTGCT 180
181 TTTCTGTGTGAATCTGAAATATGAAATCTCCAAAAGGTTTCAAGAAAGCTTGAAACA 240
181 TTTCTGTGTGAATCTGAAATATGAAATCTCCAAAAGGTTTCAAGAAAGCTTGAAACA 240
241 GTCCCGGCAATATTAATCTCCATATTTATCAATTCATGACCAATTTTGAAGAAACA 300
241 GTCCCGGCAATATTAATCTCCATATTTATCAATTCATGACCAATTTTGAAGAAACA 300
301 AAACACGAGCTGGAAGAAACAGGCGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
301 AAACACGAGCTGGAAGAAACAGGCGCAAGATTTCTGACTACATCTGATTTATGCTTAC 360
361 ATTGTGAGATGAGGCGAGATGTTAAGTACAGCAACAGCAACATGAGACGTGGAATCC 420
361 ATTGTGAGATGAGGCGAGATGTTAAGTACAGCAACAGCAACATGAGACGTGGAATCC 420
421 GTGAGACCTTATTTGCGCATATACAGAAATGATCCCGAGCTTTTCAAAAAGGCGCATCA 480
421 GTGAGACCTTATTTGCGCATATACAGAAATGATCCCGAGCTTTTCAAAAAGGCGCATCA 480
481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCAATTAAGA 540
481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCAATTAAGA 540

481 GAGCTTGTAAAGCAGTGGCGCTTGATACAAAGCAGAAATTAAGAGGTGCAATTAAGA 540
541 GGCAATCGAACAAATCGCAATTTCCCAATAGCAATGATGCTATATTTACGGCAAG 600
541 GGCAATCGAACAAATCGCAATTTCCCAATAGCAATGATGCTATATTTACGGCAAG 600
601 CCTGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
601 CCTGAGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
661 AGCAGTACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
661 AGCAGTACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 ACAGTTCGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 ACAGTTCGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 GCATTTGGACCGACCAATTAATGACCAATGATGATGATGATGATGATGATGATGATGAT 840
781 GCATTTGGACCGACCAATTAATGACCAATGATGATGATGATGATGATGATGATGATGAT 840
841 TCCGATATTAGAAAGCGCTCCACTAATTAAGAAATGCGCTCAGGCGCAATCTAGCTTC 900
841 TCCGATATTAGAAAGCGCTCCACTAATTAAGAAATGCGCTCAGGCGCAATCTAGCTTC 900
901 CAATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
901 CAATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
961 TTCAAGCAAGCAATACAGTGTGTCGAGAAATTCATCAAACTCTGTGGAGCAGCAATG 1020
961 TTCAAGCAAGCAATACAGTGTGTCGAGAAATTCATCAAACTCTGTGGAGCAGCAATG 1020
1021 AATGGGCTTACACAAAGATTTCCAGAAATGATGATGATGATGATGATGATGATGATGAT 1080
1021 AATGGGCTTACACAAAGATTTCCAGAAATGATGATGATGATGATGATGATGATGATGAT 1080
1081 ACGATCTTTTGTGCTGCGGGAATGAAAGCCGAAACCGGCAACATCATGACACAGGC 1140
1081 ACGATCTTTTGTGCTGCGGGAATGAAAGCCGAAACCGGCAACATCATGACACAGGC 1140
1141 AAGTAAAGAAATGCAATTAAGTGGAGCTACGAGAAACCTCCGCAAGCTTTGGGCT 1200
1141 AAGTAAAGAAATGCAATTAAGTGGAGCTACGAGAAACCTCCGCAAGCTTTGGGCT 1200
1201 TATGCGACCAATTAATCAACATGATGACAGTTCCTTCACTGACGACCAAGAGATGGA 1260
1201 TATGCGACCAATTAATCAACATGATGACAGTTCCTTCACTGACGACCAAGAGATGGA 1260
1261 CGGATCAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261 CGGATCAAAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1321 GCAACGGAATCTCTCTGCGGCAACCATGACAGTAAATGATGATGATGATGATGATGATGAT 1380
1321 GCAACGGAATCTCTCTGCGGCAACCATGACAGTAAATGATGATGATGATGATGATGATGAT 1380
1381 TCCATGCTACACCGATGCTGTGGAACCGTGGACAGCTTCGTGAGCATTTTGTGAAA 1440
1381 TCCATGCTACACCGATGCTGTGGAACCGTGGACAGCTTCGTGAGCATTTTGTGAAA 1440
1441 AACAGAGCATCAACCAAAACCTTCTATTAAGCGGACATGATGATGATGATGATGATGAT 1500
1441 AACAGAGCATCAACCAAAACCTTCTATTAAGCGGACATGATGATGATGATGATGATGAT 1500
1501 GACATCGGCTTTGCTACCCGAAACGATTAACCAAGATGAGGACGATGATGATGATGAT 1560
1501 GACATCGGCTTTGCTACCCGAAACGATTAACCAAGATGAGGACGATGATGATGATGATGAT 1560
1561 TCCCTGAACGTTGCTATGTAAGCAAGTCCAGTTCTATTCACACCAAAAGCGACG 1620
1561 TCCCTGAACGTTGCTATGTAAGCAAGTCCAGTTCTATTCACACCAAAAGCGACG 1620

QY 1621 TACTGTTTACTGCTGCGGCAAGCTTTGAAAACTCCCTGTAATGCTGATGCC 1680
Db 1621 TACTGTTTACTGCTGCGGCAAGCTTTGAAAACTCCCTGTAATGCTGATGCC 1680
QY 1681 CTTGCGGCAACACTGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATGCC 1740
Db 1681 CTTGCGGCAACACTGCTTCCGTAACGCTTGTCAATGATCTGACCTTGTCAATGCC 1740
QY 1741 CCAATGCGACACAGTATGTAAGAAATGACTTTTCTGCGATCAATGATTAATCGGAT 1800
Db 1741 CCAATGCGACACAGTATGTAAGAAATGACTTTTCTGCGATCAATGATTAATCGGAT 1800
QY 1801 GGGCGCAATACGTAGAAAAATGATTTATTAATGCAACCAAGCGGACGTATACAT 1860
Db 1801 GGGCGCAATACGTAGAAAAATGATTTATTAATGCAACCAAGCGGACGTATACAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGTAAT 1920
Db 1861 GAGGTACAGGCTTATTAACGTACCGGTTGACCAAGACCTTCTGTTGGCAATTTGTAAT 1920
QY 1921 TAA 1923
Db 1921 TAA 1923

RESULT 4
US-10-820-712A-2
Sequence 2, Application US/10820712A
Publication No. US20050026804A1
GENERAL INFORMATION:
APPLICANT: KAO CORPORATION
APPLICANT: Okuda, Mitsuyoshi
APPLICANT: Izawa, Yoshihumi
APPLICANT: Kobayashi, Toru
APPLICANT: Koyama, Shingo
APPLICANT: Sato, Tetsuyoshi
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 251701-USO
CURRENT APPLICATION NUMBER: US/10/820,712A
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 2003-106708
PRIOR FILING DATE: 2003-04-10
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 1923
TYPE: DNA
ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: (1)..(618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619)..()
US-10-820-712A-2

Query Match 100.0%; Score 1923; DB 8; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGACGCAATTTTGCACT 60
Db 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCAGCTGACGCAATTTTGCACT 60
QY 61 GTTGCCTTAAGTATCATCTGCAAGGTGGTGCAGAGAAATTTGATCTGCAATTCAGGA 120
Db 61 GTTGCCTTAAGTATCATCTGCAAGGTGGTGCAGAGAAATTTGATCTGCAATTCAGGA 120
QY 121 ATTCAGCAACAACTGATGCTAAAGTTTCTCAAGAGGGGCAAGCTGGTGTCTGCT 180

Db 121 ATTCAGCAACAACTGATGCTAAAGTTTCTCAAGAGGGGCAAGCTGGTGTCTGCT 180
QY 181 TTTTCTGTTGAATCTGAAAATGTAATCTCCCAAAAGTTTGCAAGAACTTGAACA 240
Db 181 TTTTCTGTTGAATCTGAAAATGTAATCTCCCAAAAGTTTGCAAGAACTTGAACA 240
QY 241 GTCCCGCAATATAAATCTCAATATATCAATTTCAATGGAACCAATTTTGAAGAAACA 300
Db 241 GTCCCGCAATATAAATCTCAATATATCAATTTCAATGGAACCAATTTTGAAGAAACA 300
QY 301 AAACAGAGCTGAAAAAAGAGGCAAGATTTCTGACATACATACCTGATTAATGCTTAC 360
Db 301 AAACAGAGCTGAAAAAAGAGGCAAGATTTCTGACATACATACCTGATTAATGCTTAC 360
QY 361 ATTGTGAGATGAGAGGCGATTTTAAGTCAAGCAACAGCACTTGAAGCACTGGAATCC 420
Db 361 ATTGTGAGATGAGAGGCGATTTTAAGTCAAGCAACAGCACTTGAAGCACTGGAATCC 420
QY 421 GTGAGGCTTAATTTGCGGATATACAGATGATCCGAGCTTTTCAAAAAGGGGCAATCA 480
Db 421 GTGAGGCTTAATTTGCGGATATACAGATGATCCGAGCTTTTCAAAAAGGGGCAATCA 480
QY 481 GAGCTGTAAAGCAAGTGGCGCTTGATACAAAGCAAAAAATTAAGAGTGCATTAAGA 540
Db 481 GAGCTGTAAAGCAAGTGGCGCTTGATACAAAGCAAAAAATTAAGAGTGCATTAAGA 540
QY 541 GGCATCGAACAAATTCGCAATTCGCAATTAAGCAATGATGTGCTATTAATTAACGGCAAG 600
Db 541 GGCATCGAACAAATTCGCAATTCGCAATTAAGCAATGATGTGCTATTAATTAACGGCAAG 600
QY 601 CTTGAGATTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 CTTGAGATTAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 AGCAGCTACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 661 AGCAGCTACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACAGGTGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACAGGTGCAATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTTGGAACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 GCATTTGGAACGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGTAATGAGAAACGCTCCATTAATTAAGAAATGCGCTCAAGGCGAATCTAGTCTTC 900
Db 841 TCCGTAATGAGAAACGCTCCATTAATTAAGAAATGCGCTCAAGGCGAATCTAGTCTTC 900
QY 901 CAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901 CAATCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TTGAGCGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TTGAGCGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 AATGGGCTTACACACAGATTCAGAAATGTAAGTATGATGATGATGATGATGATGATGAT 1080
Db 1021 AATGGGCTTACACACAGATTCAGAAATGTAAGTATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACGATCTTTTGTGCTGCGGGAATGAAGACCGAAGCGGGAACATCATGATGACCAAGGC 1140
Db 1081 ACGATCTTTTGTGCTGCGGGAATGAAGACCGAAGCGGGAACATCATGATGACCAAGGC 1140
QY 1141 ACAGCTAAATATGCAATTAACGTGAGCTACGGAACCTCGCCCAAGCTTTGGGTCT 1200
Db 1141 ACAGCTAAATATGCAATTAACGTGAGCTACGGAACCTCGCCCAAGCTTTGGGTCT 1200
QY 1201 TATGGGCAATATGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260

Db 1201 TATGCBACATATCAACATGTCGACAGTTCTTCACTGAGACCGAACAAAGATGGA 1260
Qy 1261 CGGATCAACCGGATGTCATGACCGGAAAGCTTATATCAACAAAGATCTTCTTT 1320
Db 1261 CGGATCAACCGGATGTCATGACCGGAAAGCTTATATCAACAAAGATCTTCTTT 1320
Qy 1321 GCAACGGATCTCTCTTCTGCGGCAACATGACATTAATATGATCACTATGCGTGAACG 1380
Db 1321 GCAACGGATCTCTCTTCTGCGGCAACATGACATTAATATGATCACTATGCGTGAACG 1380
Qy 1381 TCCATGCGTCAACCGATGCTGTCGGAACGTCGACAGCTTCGTAGCATTTTGTGAAA 1440
Db 1381 TCCATGCGTCAACCGATGCTGTCGGAACGTCGACAGCTTCGTAGCATTTTGTGAAA 1440
Qy 1441 AACAGAGCATCAACCAAGCTTCTATTAATAAGCGCATGATTCGCGTGAACGCT 1500
Db 1441 AACAGAGCATCAACCAAGCTTCTATTAATAAGCGCATGATTCGCGTGAACGCT 1500
Qy 1501 GACATCGGCTTGGCTACCCGAAACGGTAACCAAGATGCGGACAGTGAATTTGAATA 1560
Db 1501 GACATCGGCTTGGCTACCCGAAACGGTAACCAAGATGCGGACAGTGAATTTGAATA 1560
Qy 1561 TCCCTGAACGTTGCTATGTAACGATCAAGTCAAGTTCATCAACGCAAGCAAGAGCAGC 1620
Db 1561 TCCCTGAACGTTGCTATGTAACGATCAAGTTCATCAACGCAAGCAAGAGCAGC 1620
Qy 1621 TACTGTTTACTGCTACTGTCGCGCAAGCTTTGAAAATCTCCCTGATGCTGATGTC 1680
Db 1621 TACTGTTTACTGCTACTGTCGCGCAAGCTTTGAAAATCTCCCTGATGCTGATGTC 1680
Qy 1681 CCTGAGAGCAACATGCTCCCGTAACGCTTTCATGATCTGACCTTTCATTAACGCT 1740
Db 1681 CCTGAGAGCAACATGCTCCCGTAACGCTTTCATGATCTGACCTTTCATTAACGCT 1740
Qy 1741 CCAATGAGCAACAGTATGAGAAATGACTTATCTCCCATATCAATGATTAAGGAT 1800
Db 1741 CCAATGAGCAACAGTATGAGAAATGACTTATCTCCCATATCAATGATTAAGGAT 1800
Qy 1801 GGCCTGAATACGTAGAAAATGATTTATTAATGACCAACAAAGCGGACGTATCAATT 1860
Db 1801 GGCCTGAATACGTAGAAAATGATTTATTAATGACCAACAAAGCGGACGTATCAATT 1860
Qy 1861 GAGGTACAGGCTTATACGTACCGGTGACCAACCTTCCTGTCGATGCAATGGAAT 1920
Db 1861 GAGGTACAGGCTTATACGTACCGGTGACCAACCTTCCTGTCGATGCAATGGAAT 1920
Qy 1921 TAA 1923
Db 1921 TAA 1923

ORGANISM: Bacillus sp. KSM-KP43
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1920)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: (1) .. (618)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (619) .. ()
US-10-820-714A-2
Query Match 100.0%; Score 1923; DB 9; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGAAAAGAAAAGGTTTATCTGTTTATCAAGCTGCAGCAATTTTGTGACT 60
Db 1 ATGAGAAAAGAAAAGGTTTATCTGTTTATCAAGCTGCAGCAATTTTGTGACT 60
Qy 61 GTTGGTTAAGTAATCCATCTGACAGGTGGTGAAGAAATTTGATCTGATTTCAAAGA 120
Db 61 GTTGGTTAAGTAATCCATCTGACAGGTGGTGAAGAAATTTGATCTGATTTCAAAGA 120
Qy 121 ATTGACAAACAATGATCTTAAGTTTCTCAAGCGGCGACACTGTCGCTGCT 180
Db 121 ATTGACAAACAATGATCTTAAGTTTCTCAAGCGGCGACACTGTCGCTGCT 180
Qy 181 TTTCTGATGATCTGAAAATGTAATCTCCAAAAGGTTTTCAGAGAAAGCTTGAAACA 240
Db 181 TTTCTGATGATCTGAAAATGTAATCTCCAAAAGGTTTTCAGAGAAAGCTTGAAACA 240
Qy 241 GTTCCGCAATTAATTAATCTCCATTTATTCATTCATGACCAATTTTAAAGAAACA 300
Db 241 GTTCCGCAATTAATTAATCTCCATTTATTCATTCATGACCAATTTTAAAGAAACA 300
Qy 301 AAACAGAGCTGAAAACCAAGGCAAGATTCGACTACATGCTGATTAATGCTTAC 360
Db 301 AAACAGAGCTGAAAACCAAGGCAAGATTCGACTACATGCTGATTAATGCTTAC 360
Qy 361 ATTGTCAGTATGAGGCGCATGTTAATGTCAGCAACAGCAACATTAAGCACTGAAATCC 420
Db 361 ATTGTCAGTATGAGGCGCATGTTAATGTCAGCAACAGCAACATTAAGCACTGAAATCC 420
Qy 421 GTGAGCCTTATTTGCGCATATACAGATTAATCCCACTTTTCAACAAAAGGCGATCA 480
Db 421 GTGAGCCTTATTTGCGCATATACAGATTAATCCCACTTTTCAACAAAAGGCGATCA 480
Qy 481 GAGCTTGTAAAAGAGTGGCGCTTGATCAAGAGCAAAAATTAAGAGGTGCAATTAGA 540
Db 481 GAGCTTGTAAAAGAGTGGCGCTTGATCAAGAGCAAAAATTAAGAGGTGCAATTAGA 540
Qy 541 GGCATGAAACAATGCAATTCGCAATTAAGCAATGATGCTATATTAACGCAAG 600
Db 541 GGCATGAAACAATGCAATTCGCAATTAAGCAATGATGCTATATTAACGCAAG 600
Qy 601 CCTGATTAAGTATGATGATGATGTTGGCGGTGAAATTTGTCAAAAGCGGATGTGCG 660
Db 601 CCTGATTAAGTATGATGATGATGTTGGCGGTGAAATTTGTCAAAAGCGGATGTGCG 660
Qy 661 AGCAGTACGGGTTGATGACAAAGCAAGATGATGCGGTTGCCGATACAGGCTTGAT 720
Db 661 AGCAGTACGGGTTGATGACAAAGCAAGATGATGCGGTTGCCGATACAGGCTTGAT 720
Qy 721 ACAAGTGCATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ACAAGTGCATGACAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 781 GCAATGGAGCGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 GCAATGGAGCGACCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840
Qy 841 TCCGATTAAGAAAGGCTCACTAATAAGAAATGCGCGCTTCAGGCAATCTAGCTTC 900

Db 841 TCCGATTTAGAGAAACGGCTCCACTAATAAGAAATGCGCGCTCAGGGGAATCTAGCTTC 900
QY 901 CAATCTATCATGATATGCGGTGGGGGACCTTGAGAGACTACCTTGCAATCTGCAAACTTAA 960
Db 901 CAATCTATCATGATATGCGGTGGGGGACCTTGAGAGACTACCTTGCAATCTGCAAACTTAA 960
QY 961 TTGAGCGAAGCATATGCTGCTGGGCGAAGAAATCTATCAAACTCTGGGGAGAGAGAGT 1020
Db 961 TTGAGCGAAGCATATGCTGCTGGGCGAAGAAATCTATCAAACTCTGGGGAGAGAGAGT 1020
QY 1021 AATGGGCTTATCAACAAGATTCAGAAATGAGATGATATGTCGCAAAATGATATG 1080
Db 1021 AATGGGCTTATCAACAAGATTCAGAAATGAGATGATATGTCGCAAAATGATATG 1080
QY 1081 ACATATCTTTTCTGCTGGCGGAAATGAAGAACCGAAGCGGGAACATCATGTCACCAAGC 1140
Db 1081 ACATATCTTTTCTGCTGGCGGAAATGAAGAACCGAAGCGGGAACATCATGTCACCAAGC 1140
QY 1141 ACAGCTAAAAATGCAATTAAGTCGAGCTACCGGAAACCTCCGCGCAAGCTTTGGTCT 1200
Db 1141 ACAGCTAAAAATGCAATTAAGTCGAGCTACCGGAAACCTCCGCGCAAGCTTTGGTCT 1200
QY 1201 TATGCGGACAAATATCAACAAGATTCAGAAATGAGATGATATGTCGCAAAATGATATG 1260
Db 1201 TATGCGGACAAATATCAACAAGATTCAGAAATGAGATGATATGTCGCAAAATGATATG 1260
QY 1261 CGGATCAAAACCGGATGTCATGAGCAACCGGAAACCTTCATCATCAAGAAATCTTCTCT 1320
Db 1261 CGGATCAAAACCGGATGTCATGAGCAACCGGAAACCTTCATCATCAAGAAATCTTCTCT 1320
QY 1321 GCAACCGGATTCCTCTCTTCTGGGGGAAACATGACAGTAAATATGATATGAGGTGGAA 1380
Db 1321 GCAACCGGATTCCTCTCTTCTGGGGGAAACATGACAGTAAATATGATATGAGGTGGAA 1380
QY 1321 GCAACCGGATTCCTCTCTTCTGGGGGAAACATGACAGTAAATATGATATGAGGTGGAA 1380
Db 1381 TCCATGCTATCAACGATTCGTTGCGGAAAGTGGGACAGCTTGCGTGAAGATTTTGTGAA 1440
QY 1381 TCCATGCTATCAACGATTCGTTGCGGAAAGTGGGACAGCTTGCGTGAAGATTTTGTGAA 1440
Db 1381 TCCATGCTATCAACGATTCGTTGCGGAAAGTGGGACAGCTTGCGTGAAGATTTTGTGAA 1440
QY 1441 AACAGAGGATCAACCAAAAGCTTCTCTATTAAGACCGGCACTGATTCGCGTGCAGCT 1500
Db 1441 AACAGAGGATCAACCAAAAGCTTCTCTATTAAGACCGGCACTGATTCGCGTGCAGCT 1500
QY 1501 GACATGGGCTTGGCTATACCGGAAACGATACCAAGATGGGAGAGATGATATGATATTA 1560
Db 1501 GACATGGGCTTGGCTATACCGGAAACGATACCAAGATGGGAGAGATGATATGATATTA 1560
QY 1501 GACATGGGCTTGGCTATACCGGAAACGATACCAAGATGGGAGAGATGATATGATATTA 1560
Db 1561 TCCCTGAACGTTGCTATGTAAGAGTCCAGTTCTATCAACAGCCAAAGAGAGAG 1620
QY 1561 TCCCTGAACGTTGCTATGTAAGAGTCCAGTTCTATCAACAGCCAAAGAGAGAG 1620
Db 1561 TCCCTGAACGTTGCTATGTAAGAGTCCAGTTCTATCAACAGCCAAAGAGAGAG 1620
QY 1621 TACTGCTTTATGCTATGTAAGAGTCCAGTTCTATCAACAGCCAAAGAGAGAG 1680
Db 1621 TACTGCTTTATGCTATGTAAGAGTCCAGTTCTATCAACAGCCAAAGAGAGAG 1680
QY 1621 TACTGCTTTATGCTATGTAAGAGTCCAGTTCTATCAACAGCCAAAGAGAGAG 1680
Db 1681 CCTGCGAGCAACGCTCTCCGTAACGCTTGTCAAGATCTGACCTTGTCAATTAACGCT 1740
QY 1681 CCTGCGAGCAACGCTCTCCGTAACGCTTGTCAAGATCTGACCTTGTCAATTAACGCT 1740
Db 1681 CCTGCGAGCAACGCTCTCCGTAACGCTTGTCAAGATCTGACCTTGTCAATTAACGCT 1740
QY 1741 CCAAAATGAGCAACAGTATGAGAAATGACTTACTTCTGCAATCAATGATATTAACGAG 1800
Db 1741 CCAAAATGAGCAACAGTATGAGAAATGACTTACTTCTGCAATCAATGATATTAACGAG 1800
QY 1801 GGGCGCAATTAAGTATGAAATGATATTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1860
Db 1801 GGGCGCAATTAAGTATGAAATGATATTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1860
QY 1861 GAGGTACAGGCTTATTAAGTATGAAATGATATTAATGTAATGTAATGTAATGTAATGTAAT 1920
Db 1861 GAGGTACAGGCTTATTAAGTATGAAATGATATTAATGTAATGTAATGTAATGTAATGTAAT 1920
QY 1921 TAA 1923
Db 1921 TAA 1923

Db 1921 TAA 1923
RESULT 6
US-09-920-954-7
; Sequence 7, Application US/0920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKELI, KATSUSHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGIYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/509,814
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-09-920-954-7
Query Match 99.8%; Score 1919.8; DB 3; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCTGCTGAGCGATTTGTGACT 60
Db 1 ATGAGAAAGAAAGAAAGGTTTATCTGTTTATCTGCTGAGCGATTTGTGACT 60
QY 61 GTTGCGTTAATATCATCTGAGGTGGGCAAGAAATTTGATCTGATTTCAAGAA 120
Db 61 GTTGCGTTAATATCATCTGAGGTGGGCAAGAAATTTGATCTGATTTCAAGAA 120
QY 121 ATTGAGACAACAATGATATCTAAAGTTTCTCCAGCAGGGGCAAGCTGTGCTGCT 180
Db 121 ATTGAGACAACAATGATATCTAAAGTTTCTCCAGCAGGGGCAAGCTGTGCTGCT 180
QY 181 TTTTGTGAGATCTGAAATGTAAGAACTCCAAAGGTTTCCAGAGAAAGCTTGAACA 240
Db 181 TTTTGTGAGATCTGAAATGTAAGAACTCCAAAGGTTTCCAGAGAAAGCTTGAACA 240
QY 241 GTTCCGCAATTAATTAATCTCAATATTAATCAATTCATGAGCAATTTTGAAGAA 300
Db 241 GTTCCGCAATTAATTAATCTCAATATTAATCAATTCATGAGCAATTTTGAAGAA 300
QY 301 AAAACAGCACTGGAAGAAACAGGGGCAAGATTTCTGACTCAATCTGATTAATGCTTAC 360
Db 301 AAAACAGCACTGGAAGAAACAGGGGCAAGATTTCTGACTCAATCTGATTAATGCTTAC 360
QY 361 ATTGTGATTAAGAGGCGAGTGTAAAGTGAAGAAACAATGAGACATGAGACATGAG 420
Db 361 ATTGTGATTAAGAGGCGAGTGTAAAGTGAAGAAACAATGAGACATGAGACATGAG 420
QY 421 GTGAGACCTTATTTTGGCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGGATCA 480
Db 421 GTGAGACCTTATTTTGGCGATATACAGAAATGATCCCAAGCTTTTCAAAAAGGGGATCA 480

QY 481 GAGCTGTGTAAGCAAGTGGCGCTTGATACAAAGCAGAAAAATAAGAGGTGCAATTAAGA 540
DB 481 GAGCTGTGTAAGCAAGTGGCGCTTGATACAAAGCAGAAAAATAAGAGGTGCAATTAAGA 540
QY 541 GGCATGGAACAAATCGCACAATTGCGAATAGCAATGATGCTATATATTAACGGCAAG 600
DB 541 GGCATGGAACAAATCGCACAATTGCGAATAGCAATGATGCTATATATTAACGGCAAG 600
QY 601 CCTGAGTAATAAGGTGATGAAATGATGTTGGCGCGTGAATTTGCAACCGATGTGGCTCAG 660
DB 601 CCTGAGTAATAAGGTGATGAAATGATGTTGGCGCGTGAATTTGCAACCGATGTGGCTCAG 660
QY 661 AGCAGCTACGCGTGTGATGACAAAGCAGATGTAAGCGGTGCGGATACAGGGCTTGAT 720
DB 661 AGCAGCTACGCGTGTGATGACAAAGCAGATGTAAGCGGTGCGGATACAGGGCTTGAT 720
QY 721 ACAGGTGCGAATGACAGTTGATGATGAAAGCCTTTCGCGGGAATACTGCAATTATAT 780
DB 721 ACAGGTGCGAATGACAGTTGATGATGAAAGCCTTTCGCGGGAATACTGCAATTATAT 780
QY 781 GCATTGGGACGGAAGTAATGCAATGATACGAATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGGACGGAAGTAATGCAATGATACGAATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCGGTAATGAGAAACGGCTCCACTTAATAAGAAATGCGGCTCAGGCGAATCTAGCTTC 900
DB 841 TCGGTAATGAGAAACGGCTCCACTTAATAAGAAATGCGGCTCAGGCGAATCTAGCTTC 900
QY 901 CAATGATCATGATAGCGGTGGGGAATTGAGAGATCTTCGAAATCTGCAACTTTA 960
DB 901 CAATGATCATGATAGCGGTGGGGAATTGAGAGATCTTCGAAATCTGCAACTTTA 960
QY 961 TTCAGCCAAAGCATACAGTGTGGTGCAGAAATTCATACAAACTCCGCGGAGAGAGAG 1020
DB 961 TTCAGCCAAAGCATACAGTGTGGTGCAGAAATTCATACAAACTCCGCGGAGAGAGAG 1020
QY 1021 AATGGGGCTTACACACAGATTCACGAAATGTGATGATGATGATGATGATGATGATGATG 1080
DB 1021 AATGGGGCTTACACACAGATTCACGAAATGTGATGATGATGATGATGATGATGATGATG 1080
QY 1081 AGGATCTTTTGGCTGCGGGAATGAAAGCAGAACCGGGAACCATCATGTCACCAAGGC 1140
DB 1081 AGGATCTTTTGGCTGCGGGAATGAAAGCAGAACCGGGAACCATCATGTCACCAAGGC 1140
QY 1141 ACAGCTGTAATGCAATACAGTGGAGCTACGGAACCTCGCGGCAAGCTTTGGGCT 1200
DB 1141 ACAGCTGTAATGCAATACAGTGGAGCTACGGAACCTCGCGGCAAGCTTTGGGCT 1200
QY 1201 TATGGGGAATATCAACCATGTGGCAAGTTCTTTCAGCTGGAACCGAACAAAGATGGA 1260
DB 1201 TATGGGGAATATCAACCATGTGGCAAGTTCTTTCAGCTGGAACCGAACAAAGATGGA 1260
QY 1261 CGGATCAAAACCGGATGTGATGGAACCGGGAACGTTTCATCTACGCAAGATCTTCTCT 1320
DB 1261 CGGATCAAAACCGGATGTGATGGAACCGGGAACGTTTCATCTACGCAAGATCTTCTCT 1320
QY 1321 GCAACCGGATCTCTCTTCTGGGCGAACCATGACATGATTAATGATACATGAGGTGAAACG 1380
DB 1321 GCAACCGGATCTCTCTTCTGGGCGAACCATGACATGATTAATGATACATGAGGTGAAACG 1380
QY 1381 TCCATGAGCTACACCGATCTGCTGGAACGTTGCAAGCTTCTGAGCATTTTGTGAA 1440
DB 1381 TCCATGAGCTACACCGATCTGCTGGAACGTTGCAAGCTTCTGAGCATTTTGTGAA 1440
QY 1441 AACAGAGGATACACACCGATCTGCTGGAACGTTCTATTAATAAGGCGCATATGCGGTGACCT 1500
DB 1441 AACAGAGGATACACACCGATCTGCTGGAACGTTCTATTAATAAGGCGCATATGCGGTGACCT 1500
QY 1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAAAGATGGGGAACGATGACATTTGATTA 1560
DB 1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAAAGATGGGGAACGATGACATTTGATTA 1560
QY 1561 TCCCTGAAAGTTGCTATGTGAACGATTCAGTTCTTATTCACCAAGCCAAAAAGCGAGC 1620

DB 1561 TCCCTGAAAGTTGCTATGTGAACGATTCAGTTCTTATTCACCAAGCCAAAAAGCGAGC 1620
QY 1621 TACTGTTTACTGTACTACGCGGCAAGCCTTTGAAATCTCCGTGATAGTGTGATGCC 1680
DB 1621 TACTGTTTACTGTACTACGCGGCAAGCCTTTGAAATCTCCGTGATAGTGTGATGCC 1680
QY 1681 CCTGGACACAACTGCTTCCGTAAACGTTGTCATGATCTGACCTTGTCAATACCGCT 1740
DB 1681 CCTGGACACAACTGCTTCCGTAAACGTTGTCATGATCTGACCTTGTCAATACCGCT 1740
QY 1741 CCAATGGCACACAGTATGTAGAAATGATCTTACTTCCCATCAATGATTAACGGGAT 1800
DB 1741 CCAATGGCACACAGTATGTAGAAATGATCTTACTTCCCATCAATGATTAACGGGAT 1800
QY 1801 GCGCGCAATTAACGTAAGAAATGTATTTATTAATGACACAAAGCGGACGTATACAAAT 1860
DB 1801 GCGCGCAATTAACGTAAGAAATGTATTTATTAATGACACAAAGCGGACGTATACAAAT 1860
QY 1861 GAGGTACAGGCTTATTAACGTACCGGTGGAACACAGACCTTCTGTTGGCAATGTGAAT 1920
DB 1861 GAGGTACAGGCTTATTAACGTACCGGTGGAACACAGACCTTCTGTTGGCAATGTGAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923

RESULT 7
US-10-784-870-7
; Sequence 7, Application US/10784870
; Publication No. US20040142837A1
; GENERAL INFORMATION:
; APPLICANT: TAKATA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT FILING DATE: US/10/784, 870
; PRIOR FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1923)
US-10-784-870-7

Query Match 99.8%; Score 1919.8; DB 7; Length 1923;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1921; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGAAAGAAAGAAAGAGTTTATCTGTTTATCACTGACGCAATTTGTGCACT 60
DB 1 ATGAGAAAGAAAGAAAGAGTTTATCTGTTTATCACTGACGCAATTTGTGCACT 60
QY 61 GTTGGTTAAGTAATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120
DB 61 GTTGGTTAAGTAATCATCTGCAAGTGTGCAAGAAATTTGATCTGATTTCAAGGA 120

QY 121 ATTGAGCAAACTGATGCTAAAGTTTCTCCAGGAGGGGCGAGCTGGTCTGCT 180
DB 121 ATTGAGCAAACTGATGCTAAAGTTTCTCCAGGAGGGGCGAGCTGGTCTGCT 180
QY 181 TTTCTGATGATCTGAAAAATGTGAACTCCCAAAAGTTTGCAGAAAGACTTGAACA 240
DB 181 TTTCTGATGATCTGAAAAATGTGAACTCCCAAAAGTTTGCAGAAAGACTTGAACA 240
QY 241 GTCCCGCAATTAATTAATCTCATTTATCCAAATTCATGAGCAATTTTGAAGAAACA 300
DB 241 GTCCCGCAATTAATTAATCTCATTTATCCAAATTCATGAGCAATTTTGAAGAAACA 300
QY 301 AAAAGAGGCTGGAAAAAAGAGGGGCAAGATCTCGACTACATCACTGATTAATGCTTAC 360
DB 301 AAAAGAGGCTGGAAAAAAGAGGGGCAAGATCTCGACTACATCACTGATTAATGCTTAC 360
QY 361 ATTGCGATATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGAAGCACTGGAATCC 420
DB 361 ATTGCGATATGAGGGCGATGTTAAGTCAGCAACAAGCACTTGAAGCACTGGAATCC 420
QY 421 GTGAGGCTTATTTGCCGATATACAGATAGATCCCACTTTTCAAAAAGGGGCAATCA 480
DB 421 GTGAGGCTTATTTGCCGATATACAGATAGATCCCACTTTTCAAAAAGGGGCAATCA 480
QY 481 GAGCTGTAAAGAGTGGGCTTGAATCAAGAGAAAAATTAAGAGTGCATTTAAGA 540
DB 481 GAGCTGTAAAGAGTGGGCTTGAATCAAGAGAAAAATTAAGAGTGCATTTAAGA 540
QY 541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATATTAATTAACGGCAAG 600
DB 541 GGCATCGAACAATTCGCAATTCGCAATTAAGCAATGATGCTATATTAATTAACGGCAAG 600
QY 601 CTTGAGTATAGGTGATGAATGATGTTCCGCTGGAATTTGTCAAAAGGGATGCTCAG 660
DB 601 CTTGAGTATAGGTGATGAATGATGTTCCGCTGGAATTTGTCAAAAGGGATGCTCAG 660
QY 661 AGCAGCTACGGGTTGATGATGCAAGAGAGATGCTGAGGCTTCCGATACAGGCTTGAAT 720
DB 661 AGCAGCTACGGGTTGATGATGCAAGAGAGATGCTGAGGCTTCCGATACAGGCTTGAAT 720
QY 721 ACAGCTCGCAATGACATTCGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ACAGCTCGCAATGACATTCGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCATTGGAGCGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCATTGGAGCGAGCAATTAATGCAATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCCGATATGAGAAACGGCTCCATTAATTAAGAAATGCGCTCAGGAGCAATCTAGTCTTC 900
DB 841 TCCGATATGAGAAACGGCTCCATTAATTAAGAAATGCGCTCAGGAGCAATCTAGTCTTC 900
QY 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGAAGAACTTCCGATCTGCAAACTTGA 960
DB 901 CAATCTATCATGATAGCGGTGGGGAAGCTTGAAGAACTTCCGATCTGCAAACTTGA 960
QY 961 TTGAGCCAAAGCATACAGTGTGCTGATGCAAGATTCATCAAACTCTGAGGAGCAAGCTG 1020
DB 961 TTGAGCCAAAGCATACAGTGTGCTGATGCAAGATTCATCAAACTCTGAGGAGCAAGCTG 1020
QY 1021 AATGGGGCTTACCAACAGATTCGCAAAATGTGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AATGGGGCTTACCAACAGATTCGCAAAATGTGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ACAGTCTTTTCTGCTGCGGGAATGAAAGACCGAGCGGGAACATCACTGACCAAGCT 1140
DB 1081 ACAGTCTTTTCTGCTGCGGGAATGAAAGACCGAGCGGGAACATCACTGACCAAGCT 1140
QY 1141 ACAGTCTTTTCTGCTGCGGGAATGAAAGACCGAGCGGGAACATCACTGACCAAGCT 1200
DB 1141 ACAGTCTTTTCTGCTGCGGGAATGAAAGACCGAGCGGGAACATCACTGACCAAGCT 1200

QY 1201 TATGCGGACAAATATCAACGATGAGCAAGTCTTCACTGAGGAGCCGAAAGGATGGA 1260
DB 1201 TATGCGGACAAATATCAACGATGAGCAAGTCTTCACTGAGGAGCCGAAAGGATGGA 1260
QY 1261 CGGATCAAAACCGATGATGATGAGCAAGGAGGAGCTTCACTATCAGCAAGATCTTCTT 1320
DB 1261 CGGATCAAAACCGATGATGATGAGCAAGGAGGAGCTTCACTATCAGCAAGATCTTCTT 1320
QY 1321 GCACCGGATCTCTCTTCTGAGGAGCAACATGACAGTAAATATGATGATGATGATGATGAT 1380
DB 1321 GCACCGGATCTCTCTTCTGAGGAGCAACATGACAGTAAATATGATGATGATGATGATGAT 1380
QY 1381 TCCATGCGCTACACGATGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 TCCATGCGCTACACGATGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 AACAGAGGATCAACACCAAGCTTCTCTATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 AACAGAGGATCAACACCAAGCTTCTCTATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAGAGATGAGGAGGAGGAGGAGGAGGAG 1560
DB 1501 GACATGCGCTTGGCTACCCGAAACGGTAAACAGAGATGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 TCCCTGAAAGCTTGGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1561 TCCCTGAAAGCTTGGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
QY 1621 TACTGCTTACTCTCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
DB 1621 TACTGCTTACTCTCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
QY 1681 CTTGCGAGCAACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
DB 1681 CTTGCGAGCAACCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
QY 1741 CCAATGGGCAACAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
DB 1741 CCAATGGGCAACAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
QY 1801 GGCAGCAATACGTAAGAAATGATTAATTAATGAGCAACCAAGGAGGAGGAGGAGGAGGAG 1860
DB 1801 GGCAGCAATACGTAAGAAATGATTAATTAATGAGCAACCAAGGAGGAGGAGGAGGAGGAG 1860
QY 1861 GAGGTAACGCTTATTAACGTAACGCTTGAAGCAACAGACCTTCTGTTGGCAATGGAAT 1920
DB 1861 GAGGTAACGCTTATTAACGTAACGCTTGAAGCAACAGACCTTCTGTTGGCAATGGAAT 1920
QY 1921 TAA 1923
DB 1921 TAA 1923
RESULT 8
US-09-920-954-3
; Sequence 3, Application US/0920954
; Publication No. US20020064854A1
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKAKI, KATSUHIKA
; APPLICANT: KOBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/920,954
; PRIOR APPLICATION NUMBER: 2001-08-03
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/J998/04528

PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-09-920-954-3

Query Match 76.8%; Score 1477.2; DB 3; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 161; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 10 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGATTTTGTGACTGTTGCTTA 69
DB 7 AAGAAAGAGTGTGTTTATCTGTTTATCAGCTGACGATTTTGTGACTGTTGCTTA 66
QY 70 AGTAATCCATCTGACAGTGTGCAAGAAATTTGATCTGATTTCAAGAAATTCAGACA 129
DB 67 AACCAATCCCTCGGCTGTGATGCAAGACCTTTGATCTGATTTTAAAGAAATTCAGACA 126
QY 130 ACAATGATGTCTTAAAGTCTTCTCAAGCAGGCGACAGTGTGCTGCTTTCTGTG 189
DB 127 ACAACGATGTCTGATGTTTCTTCCAAACAGCAGCAAGTGTGCTGATTTCTGTG 186
QY 190 GAATCTGAAATGTGAACTCCCAAAAGTTTGCAAGAAAGCTTGAACAGTCCGCGCA 249
DB 187 GAGTCTGAAATGTGAACTTCTTAAGAGTTGCTTAAGAACTTGAACAGTCCGCGCA 246
QY 250 AATAATTAATCTGATTTTATCTCAATTCATGCAACCAATTTTGAAGAAACAAACAGCAG 309
DB 247 AATAATTAATCTGATTTTATCTCAATTCATGCAACCAATTTTGAAGAAACAAACAGCAG 306
QY 310 CTGGAAGAAACAGGCGCAAGATTCCTGACTACATCTGATTTATGCTTATGCTGAG 369
DB 307 CTAGAGCAATCTGAGCAAGAAATTCCTGACTACATCTGATTTATGCTTATGCTGAG 366
QY 370 TATGAGGCGATGTTAAGTCAAGCAACAGCAACATTTGAGCACTGAGATCCGTGAGCCT 429
DB 367 TATGAGGCGATGTTAAGTCAAGCAACAGCAACATTTGAGCACTGAGATCCGTGAGCCT 426
QY 430 TATTTGCCATATACAGATATAGTCCCGAGCTTTTCAAAAGGCGCATCAGACTTTGA 489
DB 427 TACTTGCCGAAATACAAATATAGTCCCGAGCTTTTCAAAAGGCGCATCAGACTTTGA 486
QY 490 AAGACAGTGGCGCTGATCAAAAGCAGAAATTAAGAGTGCATTAAGAGCATCGCA 549
DB 487 AAGACAGTGGCGCTGATCAAAAGCAGAAATTAAGAGTGCATTAAGAGCATCGCA 546
QY 550 CAAATCGACAAATTCGCAATTAAGCAATGATGCTATATATTAACGCAAGCCTGATAT 609
DB 547 GAAATGCTCAGATACGATACAGCAATGATGCTATATATTAACGCAAGCCTGATAT 606
QY 610 AAGGATGATATGATGTTGCGGTGATTTGCAAAAGCGGATGTGCTCAGAGCACTAC 669
DB 607 AAGGATGATATGATGTTGCGGTGATTTGCAAAAGCGGATGTGCTCAGAGCACTAC 666
QY 670 GGGTGTATGACAAAGACAGATCTGAGCGGTGCGGATACAGAGGCTTGAATACAGGTGCG 729
DB 667 GGGTGTATGACAAAGCGGATTTGCGAGTTGCGGATACAGAGGCTTGAATACAGAGG 726
QY 730 AATGACAGTTGATGATGATGCTTCCGCGGAAATTAATGCAATTAATGATGATGGA 789
DB 727 AACGACAGTTGATGATGATGCTTCCGCGGAAATTAATGCAATTAATGATGATGGA 786
QY 790 CGGACGATATATGCGCAATGATACGAATGCTCATGTCAGATGTGCGTCCGATATTA 849
DB 787 CGGACGATATATGCGCAATGATACGAATGCTCATGTCAGATGTGCGTCCGATATTA 846

QY 850 GGAAGCGGTCTCACTAATTAAGGAATGCGGCTCAGGGAATCTAGTCTTCAATCTATC 909
DB 847 GGAAGCGGTCTCACTAATTAAGGAATGCGGCTCAGGGAATCTAGTCTTCAATCTATC 906
QY 910 ATGATTAACCGGTGGGAGCTTGAAGACTACCTTGAATCTGCAACCTTATTCAGCCA 969
DB 907 ATGATTAACCGGTGGGAGCTTGAAGACTACCTTGAATCTGCAACCTTATTCAGCCA 966
QY 970 GATACAGTGTGCTGCTGCAAGATTTATCAAACTCTGCGGAGCAGCAGTGAATGGGCT 1029
DB 967 GCATTCAGTGTGCTGCTGCAAGATTTATCAAACTCTGCGGAGCAGCAGTGAATGGGCT 1026
QY 1030 TACACAAAGATTCAGAAATGTGATGATCTATGTCGCAAAATGATATGACATCTT 1089
DB 1027 TACACAAAGATTCAGAAATGTGATGATCTATGTCGCAAAATGATATGACATCTT 1086
QY 1090 TTGCTGCGGGAATGAAGACCGAACCGCGGAACCATCAGTGCACAGGCAAGCTTAA 1149
DB 1087 TTGCTGCGGGAATGAAGACCGAACCGCGGAACCATCAGTGCACAGGCAAGCTTAA 1146
QY 1150 AATGCAATTAACAGTGTGAGCTACGGAACCTCCGCGCACTTTGGGTCTTATGCGGAC 1209
DB 1147 AACGCAATTAACAGTGTGAGCTACGGAACCTCCGCGCACTTTGGGTCTTATGCGGAC 1206
QY 1210 AATATCAACAGATGTGCAAGTCTTCACTGATGACAGCAAGATGCAAGGATCAAA 1269
DB 1207 AATATTAACAGATGTGCAAGTCTTCACTGATGACAGCAAGATGCAAGGATCAAA 1266
QY 1270 CCGATGTCTATGACACCGGGAACGTTCACTATCAGCAAGATCTTCTTTCAGACCGGAT 1339
DB 1267 CCGATGTCTATGACACCGGGAACGTTCACTATCAGCAAGATCTTCTTTCAGACCGGAT 1336
QY 1330 TCTCTCTTCTGGGCGAACATGACAGTAAATATGATATCATGCGTGAAGCTCCATGCT 1389
DB 1327 TCTCTCTTCTGGGCGAACATGACAGTAAATATGATATCATGCGTGAAGCTCCATGCT 1386
QY 1390 AACCGATGCTGTGCGAAAGTGTGCAAGCTGTGCGAGCAATTTTGTGAAGAAACAGAGC 1449
DB 1387 AACCGATGCTGTGCGAAAGTGTGCAAGCTGTGCGAGCAATTTTGTGAAGAAACAGAGC 1446
QY 1450 ATCAACCAAGACCTTCTCTATTAAGCGGCACTGATTTGCGGATGACATGCGGCT 1509
DB 1447 ATCACTCTTAAGCCTTCTCTATTAAGCGGCACTGATTTGCGGATGACATGCGGCT 1506
QY 1510 CTGCGCTACCGGAAACGTAACCAAGATGCGGAGAGTGAATGATTAATCCCTGAAC 1569
DB 1507 CTGCGCTACCGGAAACGTAACCAAGATGCGGAGAGTGAATGATTAATCCCTGAAC 1566
QY 1570 GTTGCCTATGGAACGAGTCCAGTCTCTATTCACAGCCCAAAAGCAAGTACCTGTT 1629
DB 1567 GTTGCCTATGGAACGAGTCCAGTCTCTATTCACAGCCCAAAAGCAAGTACCTGTT 1626
QY 1630 ACTGCTATGCGGCAAGCCTTTGAAATCTCCGTGATGATGATGCGCTGCGAGC 1689
DB 1627 ACTGCAACGCGGCGCAAGCCTTTGAAATCTCCGTGATGATGATGCGCTGCGAGC 1686
QY 1690 ACAATGCTTCCGTAAGCTTGTCAATGATGGAACCTTGTCAATTAACGCTCCAAATGCG 1749
DB 1687 ACTACTGCTTGTGAACCTGCGTCAATGATGGAATTTGTCATTAACGACCAACAGGA 1746
QY 1750 ACACAGTATGTAAGAAATGACTTACTTCCGCAATACATGATTAAGTGAATGCGGCAAT 1809
DB 1747 ACACAGTATGTAAGAAATGACTTACTTCCGCAATACATGATTAAGTGAATGCGGCAAT 1806
QY 1810 AACGTAGAAATGATATTAATTAATGACCAACAAAGCGGAGCTATACATTAAGTGAACG 1869
DB 1807 AACGTAGAAATGATATTAATTAATTAATGACCAACAAAGCGGAGCTATACATTAAGTGAACG 1866
QY 1870 GCTTATTAACGTAACGCTTGTGACCAAGACCTTCTGCTGTCGATTTGATTAATTA 1923
DB 1867 GCTTATTAACGTAACGCTTGTGACCAAGACCTTCTGCTGTCGATTTGATTAATTA 1920

RESULT 9
US-10-784-870-3
Sequence 3, Application US/10784870
Publication No. US20040142837A1
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUHIKA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/10/784,870
CURRENT FILING DATE: 2004-02-24
PRIOR APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1920
TYPE: DNA
ORGANISM: *Bacillus* sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1920)
US-10-784-870-3

Query Match 76.8%; Score 1477.2; DB 7; Length 1920;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 10 AAGAAAAGGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACTGTGCTGTTA 69
DB 7 AAGAAAGGTTTATCTGTTTATCAGCTGCAGCGATTTTGTGACTGTGCTGTTA 66
QY 70 AGTATTCATCTGCGAGTGTGCAAGGATTTTGTGCTGATTTTCAAGGATTCAGACA 129
DB 67 AACCAATCCCTGCGAGTGTGCAAGGATTTTGTGCTGATTTTCAAGGATTCAGACA 126
QY 130 ACAATGATGCTAAAGGTTTCTCAAGCAGGGGCAAGCTGTGCTGCTTTTCTGTG 189
DB 127 ACAACGATGTGAGTGTGCTCAAGCAGGGGCAAGGCTGTGCTGCTTTTCTGTG 186
QY 190 GAATGTGAAGATGTGAACCTCCCAAGGTTTGCAGAGAGAGCTTGAACAGTCCGCGA 249
DB 187 GAGTCTGAAGATGTGAACCTTAAAGGATTTGCTAAAGAACTTGAACAGTCCGCGA 246
QY 250 AATATTAAGCTCATATTTATCCATTCATGAGCAATTTTGAAGAAACAAGCAGCAG 309
DB 247 AATATTAAGCTCATATTTATCCATTCATGAGCAATTTTGAAGAAACAAGCAGCAG 306
QY 310 CTGAAAAAAGGAGGCAAGGATTCGAGCTAATACCTGATTATGCTTATGCTGAG 369
DB 307 CTGAGAGCACTGAGAGCAAGGATTCGAGCTAATACCTGATTATGCTTATGCTGAG 366
QY 370 TATGAGGCGGATGTTAAGTCAAGCAACAAGCATTGAGCAGTGAATCCGTGAGCCT 429
DB 367 TATGAGGCGGATGTTCAGTCAAGCAAGGCTGCTCATTTGAACAGTGAATCACTGAGCA 426
QY 430 TATTTGCGGATTAAGGATTAAGTCCAGCTTTTCAAAAAGGCGCATCAGAGCTTGA 489
DB 427 TACTTGCAGAAATACAAATAGATCCCGCTTTTCAAAAAGGCGCATCAGAGCTGAG 486
QY 490 AAAGCAGTGGCTGTGATCAAAAGCAAAATTAAGAGTGCATTAAGAGCATCGAA 549

DB 487 AAAGCTTGGCGCTGTGATCAAGAGCAAAATTAAGAGTGCATTAAGAGCATCGAG 546
QY 550 CAATTCGACAAATTCGCAATTAAGCAATGATGTCTATATATTACGCAAAAGCTGAGTAT 609
DB 547 GAATTCGCTCAGTACGTAGCAAGCAATGACGTCTCATATATTAACGCAAAAGCTGATAT 606
QY 610 AAGGTGAATGAATGTTTGGCGGTGAATTTTCAAAAGCGAGTGTCTCAGAGCATCAG 669
DB 607 AAGGTGAATGAATGATGTGCGAGAGGTATTTCAAAAGCGAGTGTGCAAGAGCAGCTTAC 666
QY 670 GGGTTGTATGACAAGAGCAGATCTGAGCGGTTGCGATACAGGGCTGTATCAAGTCCG 729
DB 667 GGTTTGTATGACAAGCAGCAATTTGTCAGAGTTGCGATACCTGATGTGATACAGAA 726
QY 730 AATGACGTTGATGATGATGAGCTTTCGCGGAAAATTAATGCAATTAATGATGAGGA 789
DB 727 AACGACGTTGATGATGATGAGCTTTCGCGGAAAATTAATGCAATTAATGATGAGGA 786
QY 790 CGAGCAATTAATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
DB 787 CGAGCAATTAATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 850 GGAACGCGCTCACTAATTAAGAAATGCGGCTCAGGCGAATCTAGTCTTCAATCTATC 909
DB 847 GGAATTCGCGCAACCAATTAAGAAATGCGGCTCAGGCGAATCTAGTCTTCAATCTATC 906
QY 910 ATGATGAGCGGAGGAGGAGCTTGAAGACTACCTTGCAATCTGCAACCTTATTCAGCCAA 969
DB 907 ATGATGAGCGGAGGAGGAGCTTGAAGACTACCTTGCAATCTGCAACCTTATTCAGCCAA 966
QY 970 GCATACGATGCTGTGCGCAAGATTTCAACAACTCTGCGGAGCAGCAGTGAATGGGCT 1029
DB 967 GCATTCAGTGCAGGTCGCAAGATTTCAACAACTCTGCGGAGCAGCAGTGAATGGGCT 1026
QY 1030 TACACAACGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1089
DB 1027 TACACAACGATTTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 1090 TTGCTGCGGAGATGAAGAGCCGAAAGCGGAGACATCAGTGCACAGGCAAGCTTAA 1149
DB 1087 TTGCTGCGGAGATGAAGAGCCGAAAGCGGAGACATCAGTGCACAGGCAAGCTTAA 1146
QY 1150 AATGCAATTAACGATGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1209
DB 1147 AATGCAATTAACGATGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1206
QY 1210 AATATCAACATGATGAGCAAGTTCCTTCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1269
DB 1207 AATATCAACATGATGAGCAAGTTCCTTCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1266
QY 1270 CCGATGTCATGAGCAGCGGAAAGCTTCAATATCATGAGAAAGTCTTCTTGCACCGAGT 1329
DB 1267 CCGATGTCATGAGCAGCGGAAAGCTTCAATATCATGAGAAAGTCTTCTTGCACCGAGT 1326
QY 1330 TCTCTCTTCTGCGGCAACATGAGCAGTGAATTAATGATGATGATGATGATGATGATGATGAT 1389
DB 1327 TCTCTCTTCTGCGGCAACATGAGCAGTGAATTAATGATGATGATGATGATGATGATGATGAT 1386
QY 1390 AACAGGATGTTTGCAGAAAGTGTGCAAGCTTGTGAGCAATTTTGAAGAAACAAGAGC 1449
DB 1387 AACAGGATGTTTGCAGAAAGTGTGCAAGCTTGTGAGCAATTTTGAAGAAACAAGAGC 1446
QY 1450 ATCAACCAAAAGCTTCTTCAATTAAGCAGGCACTGATTCGCGTGCAGCTGATCAGTCCG 1509
DB 1447 ATCACTCCAAAGCTTCTTCAATTAAGCAGGCACTGATTCGCGTGCAGCTGATCAGTCCG 1506
QY 1510 CTGAGCTACCGGAGCGTGAACAGAGATGAGGAGCAGTGAATTAATGATGATGATGATGAT 1569
DB 1507 TTGGGTATTCGGAACGAGAAACCAAGAGATGAGGAGCAGTGAATTAATGATGATGATGAT 1566
QY 1570 GTTGCCTATGTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1629
DB 1567 GTTGCCTATGTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626

QY 1630 ACTGCTACTGCGCGAAGCCTTTGAAAATCTCCCTGGTAAGTGTGATGCCCCCTGCGAGC 1689
DB 1627 ACTGCAACGGCGGGGCAAGCATTTGAAAATCTCCCTGGTAAGTGTGATGCCCCCTGCGAGC 1686
QY 1630 ACNATGCTTCCTGTAACGCTTGTCAATGATCTGGAACCTTGTCATTAACGCTCCCAATAGC 1749
DB 1687 ACTATGCTTCCTGTAACGCTTGTCAATGATCTGGAACCTTGTCATTAACGCTCCCAATAGC 1746
QY 1750 ACACGCTATGTAAGAAATGATCTTCTGCAATGATTAATGATGAGATGCGCGCAAT 1809
DB 1747 ACNATATGCTGGAATGATCTTCTGCAATGATTAATGATGAGATGCGCGCAAT 1806
QY 1810 AACGTGAATAATGATTTATTAATGATCAACCAAGCGGAGCTATTAATGATGAGTACAG 1869
DB 1807 AACGTGAATAATGATTTATTAATGATGCGCCCAAGGGAACATATACCTGAGGTGCA 1866
QY 1870 GCTTATTAACGTACCGGTGGAACCAAGACCTTCTGTTGGCAATGATTAATTA 1923
DB 1867 GCATATTAATGTCGCGGTGGAACCAACCACTTCTGTTGGCAATGATTAATTA 1920

RESULT 10

US-10-385-662-1
/ Sequence 1, Application US/10385662
/ Publication No. US20040002432A1
/ GENERAL INFORMATION:
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SATO, TSUYOSHI
/ APPLICANT: SAITO, KAZUHIRO
/ APPLICANT: SUMITOMO, NOBUYUKI
/ APPLICANT: IZAWA, YOSHIFUMI
/ APPLICANT: SAKIKI, KATSUHIISA
/ APPLICANT: KOBAYASHI, TOHRU
/ APPLICANT: NOMURA, MASAFUMI
/ TITLE OF INVENTION: Alkaline protease
/ FILE REFERENCE: 234938US0
/ CURRENT APPLICATION NUMBER: US/10/385,662
/ CURRENT FILING DATE: 2003-03-12
/ PRIOR APPLICATION NUMBER: JP 2002-081428
/ PRIOR FILING DATE: 2002-03-22
/ PRIOR APPLICATION NUMBER: JP 2002-165987
/ PRIOR FILING DATE: 2002-06-06
/ PRIOR APPLICATION NUMBER: JP 2002-304230
/ PRIOR FILING DATE: 2002-10-18
/ PRIOR APPLICATION NUMBER: JP 2002-304231
/ PRIOR FILING DATE: 2002-10-18
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1305
/ TYPE: DNA
/ ORGANISM: Bacillus sp. KSM-KP43
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1305)
/ OTHER INFORMATION:
/ US-10-385-662-1

Query Match 67.9%; Score 1305; DB 6; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGGCTCAAGACGACTACGGGTGTAT 678
DB 1 AATGATGTTGCGCGTGAATTTGCAAGCGGATGTGGCTCAAGACGACTACGGGTGTAT 60
QY 679 GGAACAAGACATGCTATGACGCTTGGCCGATACAGGGCTTGATACAGGTGGCAATGACAT 738
DB 61 GGAACAAGACATGCTATGACGCTTGGCCGATACAGGGCTTGATACAGGTGGCAATGACAT 120
QY 739 TCGATGATGAAGCCTTCCGCGGAAATTAATGATTAATGATTAATGATTAATGATTAATGAT 798
DB 739 TCGATGATGAAGCCTTCCGCGGAAATTAATGATTAATGATTAATGATTAATGATTAATGAT 798

DB 121 TCGATGATGAAGCCTTCCGCGGAAATTAATGATTAATGATTAATGATTAATGATTAATGAT 180
QY 799 AATGCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 181 AATGCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 859 TCCGCTAATTAAGGAATGCGGCTCAAGCGGATGATGATGATGATGATGATGATGATGATGAT 918
DB 241 TCCGCTAATTAAGGAATGCGGCTCAAGCGGATGATGATGATGATGATGATGATGATGATGAT 300
QY 919 GGTGGGGGACTTGGAGGACTACCTTGGATATGCAAACTTATTCAGCGCAAGCATACAGT 978
DB 301 GGTGGGGGACTTGGAGGACTACCTTGGATATGCAAACTTATTCAGCGCAAGCATACAGT 360
QY 979 GCTGGTGCAGAAATTCATACAACTCTGGGGGACACAGTGAATGGGGCTTTACACA 1038
DB 361 GCTGGTGCAGAAATTCATACAACTCTGGGGGACACAGTGAATGGGGCTTTACACA 420
QY 1039 GATTCCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
DB 421 GATTCCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1099 GGAATGAAAGACCGAACCGCGGAAACCATGATGATGATGATGATGATGATGATGATGATGAT 1158
DB 481 GGAATGAAAGACCGAACCGCGGAAACCATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1159 AAGTGGAGCTACGGAACCTCTCGCCCAAGCTTTGGGTATTAATGGGACAAATATCAAC 1218
DB 541 AAGTGGAGCTACGGAACCTCTCGCCCAAGCTTTGGGTATTAATGGGACAAATATCAAC 600
QY 1219 CATGTGGCAGATCTCTGTCAGTGAACCGAACCGAAGATGGAACCGGATCAACCGGATGTC 1278
DB 601 CATGTGGCAGATCTCTGTCAGTGAACCGAACCGAAGATGGAACCGGATCAACCGGATGTC 660
QY 1279 ATGGACCGGGAAGGTTCAATCTATCAAGAAATCTTCTCTTGGACCGGATTCCTCTTC 1338
DB 661 ATGGACCGGGAAGGTTCAATCTATCAAGAAATCTTCTCTTGGACCGGATTCCTCTTC 720
QY 1339 TGGGCGAACCATGACATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
DB 721 TGGGCGAACCATGACATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1399 GTTGTGGGAAAGTGGACAGCTTCTGTAAGATTTTGGAAACAGAGGATGATGATGATGATGAT 1458
DB 781 GTTGTGGGAAAGTGGACAGCTTCTGTAAGATTTTGGAAACAGAGGATGATGATGATGATGAT 840
QY 1459 AAGCTTCTCTATTAAGGCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1518
DB 841 AAGCTTCTCTATTAAGGCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1519 CCGAACGTTAACCAAGATGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1578
DB 901 CCGAACGTTAACCAAGATGAGGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1579 GTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638
DB 961 GTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1639 GCCCGCAAGCTTTGAAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1698
DB 1021 GCCCGCAAGCTTTGAAAATCTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1699 TCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
DB 1081 TCCGTAAGCTTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1759 GTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
DB 1141 GTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1819 AATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1878
DB 1201 AATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260

SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-9865
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Protease 9865
US-10-820-712A-14

Query Match 66.7%; Score 1281.8; DB 9; Length 1302;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1289; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 619 AATGATGTTGCCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACCGGTTGTAT 678
DB 1 AATGATGTTGCAAGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACCGGTTGTAT 60
QY 679 GGAACAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTGCGCAATGACAGT 738
DB 61 GGAACAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTGCGCAATGACAGT 120
QY 739 TCGATGATGAAAGCTTCCGCGGGAATAATTAATGATATATGCAATTTGGAGCGAGCAAT 798
DB 121 TCGATGATGAAAGCTTCCGCGGGAATAATTAATGATATATGCAATTTGGAGCGAGCAAT 180
QY 799 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
DB 181 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 859 TCCACTAATTAAGGAATGGCGGCTCAGCGGAATCTAGCTTCATCTCATATGATGATGATGAT 918
DB 241 TCCACTAATTAAGGAATGGCGGCTCAGCGGAATCTAGCTTCATCTCATATGATGATGATGATGAT 300
QY 919 GGTGGGGAAGCTTGAAGAGCTACCTTTCGAACTGCAAACTTTATTTGACGCAAGCATACAGT 978
DB 301 GGTGGGGAAGCTTGAAGAGCTACCTTTCGAACTGCAAACTTTATTTGACGCAAGCATACAGT 360
QY 979 GGTGGGGAAGCTTGAAGAGCTACCTTTCGAACTGCAAACTTTATTTGACGCAAGCATACAGT 1038
DB 361 GGTGGGGAAGCTTGAAGAGCTACCTTTCGAACTGCAAACTTTATTTGACGCAAGCATACAGT 420
QY 1039 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
DB 421 GATTCCAGAAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 1099 GGGAAATGAAAGACCGAAGCGGGAACCATGATGATGATGATGATGATGATGATGATGATGAT 1158
DB 481 GGGAAATGAAAGACCGAAGCGGGAACCATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1159 ACAGTCGAGAGCTACGGAAGAACTCCGCCAAGCTTTGGGTCTTATGCGGACCAATATTCAC 1218
DB 541 ACAGTCGAGAGCTACGGAAGAACTCCGCCAAGCTTTGGGTCTTATGCGGACCAATATTCAC 600
QY 1219 CATGTGGAACAGTTCTCTTCAAGTGAAGCGAACAAGATGATGATGATGATGATGATGATGATGAT 1278
DB 601 CATGTGGAACAGTTCTCTTCAAGTGAAGCGAACAAGATGATGATGATGATGATGATGATGATGAT 660
QY 1279 AATGCCAAGCGGAAGCTTCAATCTATCAGCAAGATCTTCTTTGCAACCGGATTCCTCTTC 1338
DB 661 AATGCCAAGCGGAAGCTTCAATCTATCAGCAAGATCTTCTTTGCAACCGGATTCCTCTTC 720
QY 1339 TGGGGAACCAATGACAGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
DB 721 TGGGGAACCAATGACAGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 1399 GTTGTGGAACAGTGTGCAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
DB 781 GTTGTGGAACAGTGTGCAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1459 AAGCTTCTCTATTAAGGCGCACTGATTCGCGGTGCACTGATGATGATGATGATGATGATGATGAT 1518
```

```
DB 841 AAGCTTCTCTATTAAGGCGCACTGATTCGCGGTGCAAGACATGCGCTTGGCTAC 900
QY 1519 CCGAAGGTAAACCAAGATGAGGAGCAAGTGAATTTGATTAATCCCTGAAGCTTGCTAT 1578
DB 901 CCGAAGGTAAACCAAGATGAGGAGCAAGTGAATTTGATTAATCCCTGAAGCTTGCTAT 960
QY 1579 GTGAACGATGCAAGTTCTCTATTCACACGCGCAAAAGGAGCGTACTCGTTTATGCTACT 1638
DB 961 GTGAACGATGCAAGTTCTCTATTCACACGCGCAAAAGGAGCGTACTCGTTTATGCTACT 1020
QY 1639 GCCGCAAGCTTTGAAATATCCCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1698
DB 1021 GCCGCAAGCTTTGAAATATCCCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1699 TCCGTAAGCTTTGCAATGATCTGGAAGCTTGTCAATTCAGCTTCAAAATGGAACAGATGAT 1758
DB 1081 TCCGTAAGCTTTGCAATGATCTGGAAGCTTGTCAATTCAGCTTCAAAATGGAACAGATGAT 1140
QY 1759 GTGAAGATGATCTTACTTCCGCAATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
DB 1141 GTGAAGATGATCTTACTTCCGCAATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1819 AATGATTTTAAATGACCAACAAGGCGGAGCGTATCAATGATGATGATGATGATGATGATGATGAT 1878
DB 1201 AATGATTTTAAATGACCAACAAGGCGGAGCGTATCAATGATGATGATGATGATGATGATGATGAT 1260
QY 1879 GTACCGGTTGACCAAGACCTTCTCGTTGCAATTTGAA 1919
DB 1261 GTACCGGTTGACCAAGACCTTCTCGTTGCAATTTGAA 1301

RESULT 13
US-10-820-712A-11
; Sequence 11, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuyoshi
; APPLICANT: Izawa, Yoshifumi
; APPLICANT: Kobayashi, Tohru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251701-USO
; CURRENT APPLICATION NUMBER: US/10/820, 712A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus sp. KSM-kp9860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Protease KP9860
US-10-820-712A-11

Query Match 51.8%; Score 997; DB 8; Length 1302;
Best Local Similarity 85.4%; Pred. No. 5.1e-287;
Matches 1111; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 619 AATGATGTTGCCGCGTGAATTTGCAAGCGGATGTGCTCAGAGCAGTACCGGTTGTAT 678
DB 1 AATGATGTTGCCGCGGAGATTTGGAAGCGGATGTGCTCAGAGCAGTACCGGTTGTAT 60
QY 679 GGAACAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTGCGCAATGACAGT 738
DB 61 GGAACAAGACAGATCGTAGCGGTTGCCGATACAGGGCTTGATACAGGTGCGCAATGACAGT 120
QY 739 TCGATGATGAAAGCTTCCGCGGGAATAATTAATGATATATGCAATTTGGAGCGAGCAAT 798
```

Db 121 TCGATGCAATGAAGCTTCGCGGGTAATAAAGCACTATATGCACTGGGTGGAGCAAT 180
Qy 799 AATGCCAATGATACGAATGTGATGTGATACGATGTGGCTCGGTATAGGAACGGC 858
Db 181 AATGGAAATGATACGAACGGTATGATCCCATGTGGCAAGTTCGGTATAGGAATGGC 240
Qy 859 TCCACTAATAAAGGAATGGGCTCAGGGGAATCTATGCTTCCATCTATCATGTATAGC 918
Db 241 GCATGGAATTAAGGAATGGCACTCAGAGGAATCTGGTTTTCATATCCATCATGATATAC 300
Qy 919 GGTGGGGGACTTGAAGACTACCTTGAATCTGCAACCTTATTCAGCAAGCATACAGT 978
Db 301 AGTGGGGGCTTGAAGGCTTGGCTTCCATCTGCAAACTTATTCAGCAAGCATACAGT 360
Qy 979 GCTGTGCGCAATTCATCAAACTCTGGGGAGCAGCAGTGATGGGCTTACACACA 1038
Db 361 GCAAGGTGCGCAATTCATCAAACTCTGGGGAGCAGCGGTGATGGGGCTTACACACA 420
Qy 1039 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATCTCTTTCGCTGC 1098
Db 421 GATTCAGAAATGTGATGATCTATGTAGAAAAAATGATATGATCAATTCCTTTCGCGCT 480
Qy 1099 GGGAAATGAAGACCGAACGGCGGATCATCATGTGCACTGTGATCGGCTAATAACGGCATC 1158
Db 481 GGGAAATGAAGGCGCAACGGCGGATCATCATGTGCACTGTGATCGGCTAATAACGGCATC 540
Qy 1159 ACAGTGCGAGTACGGAACCTCCGCGCCAGCTTGGGTCTTATGCGGACATATCAAC 1218
Db 541 ACTGTGCGGCGCAACGGAACCTCGCGCTCAAGCTTGGGTCTTATGCAATATATTAAC 600
Qy 1219 CATGTGCGCAATTCCTTCACTGATGACCGAAGATGAGAGATGAGAGATGAGATGAG 1278
Db 601 CAGGTGCGCAATTCCTTCACTGATGACCGAAGATGAGAGATGAGAGATGAGATGAG 660
Qy 1279 ATGGACCGGGAAAGCTTCACTATCATGAGAAATCTTCTTGGCAACGGATTCCTCTTC 1338
Db 661 ATGGACCGGGAAAGCTTCACTATCATGAGAAATCTTCTTGGCAACGGATTCCTCTTC 720
Qy 1339 TGGGCGAATCATGACAGCAATATGTGCTATGTGGGAGAACGTCCATGCAACCGATC 1398
Db 721 TGGGCGAATCATGACAGCAATATGTGCTATGTGGGAGAACGTCCATGCAACCGATC 780
Qy 1399 GTTGTGGAAGAGTGGCAAGCTTGTGAGCATTTTGTGAAAAACAGAGCATCACACA 1458
Db 781 GTTGTGGAAGAGTGGCAAGCTTGTGAGCATTTTGTGAAAAACAGAGCATCACACTCT 840
Qy 1459 AAGGCTTCTCTATTAAGCGGCACTGATTCGCGGAGCTGACATCGGCTTGGCTAC 1518
Db 841 AAGGCTTCTCTATTAAGCGGCTTGTGATTCGAGGTGCTGATGTGATGGGTATC 900
Qy 1519 CGGAACGATTAACGAATGAGGAGAGAGTGAATGATGATTAATCCCTGAACGTTGCTAT 1578
Db 901 CGGAACGATTAACGAATGAGGAGAGAGTGAATGATGATTAATCCCTGAACGTTGCTAT 960
Qy 1579 GTGAACGAGTCCAGTCTCTATTCACAGCCAAAAACGAGCATCTGTTTACTGCTACT 1638
Db 961 GTGAACGAGTCCAGTCCCTATCACTACCAAAAAACGAGCATATCTTTTACTGCAAG 1020
Qy 1639 GCGGCGAAGCTTTGAAAAATCTCCCTGTATGTGATGATGATGATGATGATGATGATG 1698
Db 1021 GCGGCGAAGCTTTGAAAAATCTCCCTGTATGTGATGATGATGATGATGATGATGATG 1080
Qy 1699 TCCCTGAACGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1758
Db 1081 TCCCTGAACGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy 1759 GTAGGAATGATCTTTTCTTGGCATACAAATGATTAATGATGATGATGATGATGATGAT 1818
Db 1141 GTAGGAATGATCTTCTTGGCATACAAATGATTAATGATGATGATGATGATGATGATGAT 1200
Qy 1819 AATGATTTATTAATGACCAAGCGGAGGATTAATGATGATGATGATGATGATGATGATGAT 1878

Db 1201 AATGATTTATTAATTCGCCCCCAAGTGAACATATACATGATGATGATGATGATGAT 1260
Qy 1879 GTACCGGTTGACACACAGCTTCTCTGTTGGCAATTTGAA 1919
Db 1261 GTCCGGTTGACACAAAACTTCTGTTGGCAATTTGAA 1301

RESULT 14
US-10-820-714A-12
; Sequence 12, Application US/10820714A
; Publication No. US20050214922A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Kobayashi, Toru
; APPLICANT: Sumitomo, Nobuyuki
; APPLICANT: Takimura, Yasushi
; APPLICANT: Sato, Teiyoushi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251697US0
; CURRENT APPLICATION NUMBER: US/10/820,714A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106709
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: *Bacillus* sp. KSM-Kp9860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: protease KP9860
US-10-820-714A-12

Query Match 51.8%; Score 997; DB 9; Length 1302;
Best Local Similarity 85.4%; Pred. No. 5,1e-287;
Matches 111; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

Qy 619 AATGATGTTGCGCGTGAATTTGTCAAGCGATGTGCTCAGAGAGCATACGGTGTAT 678
Db 1 AATGATGTTGCGCGAGATTTGTGAAGGATGTGCAACAGAGCATACGGTGTAT 60
Qy 679 GGAAGAAGACAGATGTAGCGGTTCCGATACAGGCTTGTATCAGGTGCAATGACAT 738
Db 61 GGAAGAAGACAGATTTGTGCGAGTTCCGATACGATTTGATGATGATGATGATGATGAT 120
Qy 739 TCGATGCAATGAAGCTTCCGCGGAAATTAATCTGCAATTAATGATGATGATGATGATGAT 798
Db 121 TCGATGCAATGAAGCTTCCGCGGTAATTAATGATGATGATGATGATGATGATGATGAT 180
Qy 799 AATGCCAATGATACGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
Db 181 AATGCCAATGATACGAACGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 859 TCCACTAATAAAGGAATGGGCTCAGGGGAATCTATGCTTCCATCTATCATGTATAGC 918
Db 241 GCATGGAATTAAGGAATGGCACTCAGAGGAATCTGGTTTTCATATCCATCATGATATAC 300
Qy 919 GGTGGGGGACTTGAAGACTACCTTGAATCTGCAACCTTATTCAGCAAGCATACAGT 978
Db 301 AGTGGGGGCTTGAAGGCTTGGCTTCCATCTGCAAACTTATTCAGCAAGCATACAGT 360
Qy 979 GCTGTGCGCAATTCATCAAACTCTGGGGAGCAGCAGTGATGGGCTTACACACA 1038
Db 361 GCAAGGTGCGCAATTCATCAAACTCTGGGGAGCAGCGGTGATGGGGCTTACACACA 420
Qy 1039 GATTCCAGAAATGTGATGATCTATGTGCGCAAAATGATATGATCTCTTTCGCTGC 1098
Db 421 GATTCAGAAATGTGATGATCTATGTAGAAAAAATGATATGATCAATTCCTTTCGCGCT 480
Qy 1099 GGGAAATGAAGACCGAACGGCGGATCATCATGTGACCAAGGCAAGCTAATAATGCAAT 1158

```
Db 481 GGGATGAAAGGCGGACGCGGTACATCACTGACCTGGTAAACCCATC 540
Qy 1159 ACAATCGAGACATCGGAAAACCTCCGCCAGCTTGGCTTTATGCGGACATATCAAC 1218
Db 541 ACTGTCGGGCGCAACGGAACCTGGCTCCAGCTTCGTTCTCATGCAAGATATATTAAC 600
Qy 1219 CATGTCGACAGTCTCTTCACTGAGACCGGACAAAGAGATGACGATCAACCGGATGTC 1278
Db 601 CAGCTTGACAGATCTCTCCGTCGCGCGGACAAAGAGATGAGGCGAATCAAGCTTATGTC 660
Qy 1279 ATGGACCGGGAACCTTCACTATCAAGACAGATCTTCTTTCGACCGGATTCCTCTTC 1338
Db 661 ATGGCGGACGAGACATACATTTTATCAGACAGATCTTCTTTCGACCGGATTCCTCTTC 720
Qy 1339 TGGGCGAACCATGACAGTAAATATGATATCATGGGTGGAACGTCATGCGTCAACCGATC 1398
Db 721 TGGGCGAATATGACAGCAAAATATGCTATATGGGTGGAACGTCATGCGCAACCGAAT 780
Qy 1399 GTTGTGGAACGTCGACAGCTTGTGAGCATTTTGTGAAAAGAGAGGACATCAACCA 1458
Db 781 GTTGTGCGGAAATGTGACAGCTCCGTGAGCATTTTGTGAAAAGAGAGAATCACTCT 840
Qy 1459 AAGCTTCTATATTAAGCGGACATGATGCGGTCGAGCTGACATCGGCTTGGCTAC 1518
Db 841 AAGCTTCTATATTAAGAGAGCTTGAATGCAAGGTGCTGATGTTGGATTTGGTTAT 900
Qy 1519 CCGAACGCTTAACCAAGATGCGGACGATGATGCAATTTGCAATTCCTTGAACGTTGCTAT 1578
Db 901 CCGAACGGAACCAAGATGCGGCGAGTGAACCTGGAATTAATCGTTGAACGTTGCTAT 960
Qy 1579 GTGAACGAGTCACTTCTATCAACGCGCAAAAAGGAGGATCGTTACTGTTACTGCTCT 1638
Db 961 GTGAACGATCACTGCTCTATCACTGACCAAAAAGGAGCATATACCTTTACTGCAACG 1020
Qy 1639 GCCCGCAAGCTTTGAAAATCTCCGTGATGCTGATGCTCCGTCGAGCAACATGCT 1698
Db 1021 GCGGCGAAGCATTAATAATCTCCGTGATGCTGATGCTCCGTCGAGCAACATGCTCT 1080
Qy 1699 TCCGTAAGCTTGTCAATGATGTCGACCTTGTCAATTCGCGTCAATGCGACACAGAT 1758
Db 1081 TCTGTAACTGTGTCAATGATTTGATTTGATTCATTAACGACCAAGCAAGATAT 1140
Qy 1759 GTAAGAAATGACTTACTTCCGATACATGATGATACTGGGATGCGCGAATTAAGCTA 1818
Db 1141 GTCCGGAATGACTTCTGACACCATTTGACATTAATGCGGATGCGCGAATTAAGTAA 1200
Qy 1819 AATGATTTATTAATGACACCAAGCGGACGATATCAATTAAGGTAAGGCTTATTAAC 1878
Db 1201 AATGATTTATTAATTTGCCCCCAAGTGAACATATTAACATTAAGTCAAGTCAAGTAA 1260
Qy 1879 GTACCGGTTGACCAACAGACTTCTGTTGGCAATTTGAA 1919
Db 1261 GTCCGCGTTGACCAAAACTTCTGTTGGCAATTTGAA 1301

RESULT 15
US-10-820-712A-21
; Sequence 21, Application US/10820712A
; Publication No. US20050026804A1
; GENERAL INFORMATION:
; APPLICANT: KAO CORPORATION
; APPLICANT: Okuda, Mitsuoshi
; APPLICANT: Izawa, Yoshitumi
; APPLICANT: Kobayashi, Toru
; APPLICANT: Koyama, Shingo
; APPLICANT: Sato, Tsuyoshi
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 251701-USO
; CURRENT APPLICATION NUMBER: US/10/820,712A
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 2003-106708
; PRIOR FILING DATE: 2003-04-10
```

```
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Bacillus NCIB12289
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: Gene name Al, protease Al
US-10-820-712A-21

Query Match 44.4%; Score 854.6; DB 8; Length 1302;
Best Local Similarity 76.6%; Pred. No. 2.2e-244;
Matches 1022; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

Qy 619 AATGATGTTGCGCGTGAATTTGTCAAAGCGGATGCTCAGAGACGTCAGGTTGAT 678
Db 1 AACGATGTTGCAAGGCAATTTAAAGCGATGTGCCAGAGACGTCATGTTATAT 60
Qy 679 GGAACAGACAGATGCTAGCGGTTGCGATACAGGCTTGAATACAGTTCGACATGACGT 738
Db 61 GGGCAAGGCGAAGTGTGACATGCGGATACCGGACTGATACAGGCGTAAAGACGC 120
Qy 739 TCGATGCAATGAAGCTTCCGCGGGAATTAATGATTAATGATTAATGATGGAACGAT 798
Db 121 TCGATGATGAAGCTTCCGAGAAAGATTAACCGGATATATGCTTGGAAACAAAC 180
Qy 799 AATGCAATGATCAAGATGCTCATGATGCGATGCTGCTCGTATTAAGAAACGCG 858
Db 181 AAGCCATGATCAACAAACGACGCGGACGATGTTGCGGATCGTTTGAAGAACGT 240
Qy 859 TCCACTAATTAAGAAATGCGCTCAAGGCAATTAATGCTTCCATCTATATGATGAC 918
Db 241 ACTTCGATTAAGGATGCGCTCAACAGCTAACTTAATGCTTAATGATGACACGC 300
Qy 919 GATGCGGAGCTTGAAGGACTCTTCGATGTCGAAACCTTAATCAGCGCAACATACGT 978
Db 301 AATGCGGACTTGGCGGACTGCTTCGATTAATGATTAATTAATTAATGCGGCAATAT 360
Qy 979 GCTGTGCAAGATTCATCAAACTCTGCGGAGCAAGTGAATGCGGCTTACACAA 1038
Db 361 GCGGTGCAAGATTCATCAAGATTCATGAGGAGCGCTGAAACGAGGCTACATAC 420
Qy 1039 GATTCAGAAATGATGATGATGATGCGCAAAATATATGATGATGCTTTGGCTGCC 1098
Db 421 GATTCAGAAACGTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 1099 GGAATGAAGGACGGAACGCGGGAACCATGATGACACAGGCAAGCTAAATGCAATA 1158
Db 481 GTTAACGAAGGCGGAATGCGGCAACATGATGCTCTGTATCCGGAATGCTATC 540
Qy 1159 ACAATCGAGACATCGGAAAACCTCCGCCAGCTTGGCTTTATGCGGACATATCAAC 1218
Db 541 ACAATCGAGGCAACAGAAAACCTCCGCCAGCTTGGATCTTATGCTGACATCAAT 600
Qy 1219 CATGTGACAGTCTCTTCACTGAGACCGACAAAGATGAGCGGATCAACCGGATGTC 1278
Db 601 CATGTACACAGTCTCTCCCGCGGACCTACAAAGATGAGCGGATCAACCGGATGTC 660
Qy 1279 ATGGACCGGGAACGTTCACTATCAAGACAGATCTTCTGTCGACCGGATTCCTCTTC 1338
Db 661 ATGGACCGGGAACATTTATTTATTCGCAAGATCTTCTGCTCCGACCTCTCATTC 720
Qy 1339 TGGGCGAACATGACAGTAAATATGACATATGAGTGAACGTCATGCTACACGATC 1398
Db 721 TGGCAACCATGACAGCAAAATATGCTTAATGAGTGAACATCATGCGGACCGAT 780
Qy 1399 GTTGTGGAACGTCGACAGCTTGTGAGCATTTTGTGAAAAGAGGCGATCAACCA 1458
Db 781 GTACTGTAACTGTTGACAGTTACGTTGACATTTCAAAAACAGAGGATCACTCT 840
Qy 1459 AAGCTTCTATTAAGGCGACATGATGCGGTCAGCTGACATCGGCTTGGCTAC 1518
```

Db 841 AAACCAATCTTGTGTAAGCAGCTCTATATGCCGAGCAACTGATATCGGTCTTGCTAT 900
QY 1519 CCGAAGGTAAACCAAGATGGGACGAGTACATTTGATTAATCCCTGAACGTTGCTAT 1578
Db 901 CCGAGTGGAAACCAAGATGGGAGATGACATTTGACAACTTAAATGTAGCTTTC 960
QY 1579 GTGAACGAGTCCAGTCTCTATCCACGACCAAAAAGCGAGTACTGTTACTGCTACT 1638
Db 961 GTAAATGAAACAGCTCTTATCTATACCAAAAGCTAGTATTCATTTACTGACAA 1020
QY 1639 GCCGGAAGCTTTGAAAATCTCCCTGTATGATCTGATGCCCTGCGACACAATGCT 1698
Db 1021 TCAGGCAAACTTTGAAGATTTCAATGGTTGGTCTGATGCAACGCAAGTACTTCGCA 1080
QY 1699 TCCGTAAAGCTTGTCAATGATCTGACCTTGTCAATCCGCTCCAAATGGCAACAATAT 1758
Db 1081 TCGGTACATTTGGTGAATGATCTGATCTGTGATTAACGCTCCAAATGGACAAAGTAT 1140
QY 1759 GTAGAAATGACTTTACTTCCGCAATCAATGATTAATGAGATGCGGCAATTAAGTAA 1818
Db 1141 GTTGAACGACTTTACTGCTCCCTATGATTAATTAATGAGATGAGACGTAACAATGTAAG 1200
QY 1819 AATGATTTTAAATGACACCAAAAGCGGACGTATACATTTGAGTACAGGCTTAAAC 1878
Db 1201 AAGGTGTTATCAATGCTCCGCAAAAGGGAACGTATACAGTTGAGTACAGGCTTAAAC 1260
QY 1879 GTACCGGTGACCAACAGACTTCTGTTGGCAATTTGTAA 1919
Db 1261 GTTCCACAAGGCGCGAGCGTTCCTTGGCTATTTGTAA 1301

Search completed: April 8, 2006, 07:34:29
JOD time : 1384.58 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2006, 03:01:47 ; Search time 897.7 Seconds

(without alignments)
8569.753 Million cell updates/sec

Title: US-10-784-870-5

Perfect score: 1923

Sequence: 1 atgagagaagagaagaaagt.....cgttcgcatctgtaattaa 1923

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs / 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA New:*
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB_seq.*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB_seq.*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB_seq.*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB_seq.*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB_seq.*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB_seq.*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB_seq.*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB_seq.*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB_seq.*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB_seq.*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB_seq.*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB_seq.*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB_seq.*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB_seq.*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.2	2.1	604	US-09-925-065A-3789	Sequence 3789, App
2	40.2	2.1	604	US-10-301-480-105026	Sequence 105026,
3	40.2	2.1	604	US-10-301-480-718435	Sequence 718435,
4	39.8	2.1	1996	US-09-925-065A-718697	Sequence 718697,
5	39.6	2.1	417	US-09-925-065A-562287	Sequence 562287,
6	39.6	2.1	514	US-09-925-065A-124623	Sequence 124623,
7	39.6	2.1	527	US-10-301-480-75030	Sequence 75030, A
8	39.6	2.1	527	US-10-301-480-688439	Sequence 688439,
9	39.6	2.1	528	US-10-301-480-222275	Sequence 222275,
10	39.6	2.1	528	US-10-301-480-835684	Sequence 835684,
11	37.4	1.9	655	US-09-925-065A-841440	Sequence 841440,
12	37.4	1.9	656	US-09-925-065A-31269	Sequence 31269, A
13	37.4	1.9	1766	US-09-925-065A-132507	Sequence 132507,
14	37.4	1.9	1766	US-10-301-480-745916	Sequence 745916,
15	37.4	1.9	594	US-10-301-480-1251	Sequence 1251, App
16	37.2	1.9	594	US-10-301-480-614660	Sequence 614660,
17	37.2	1.9	594	US-10-301-480-614660	Sequence 614660,
18	36.4	1.9	1713	US-10-750-185-54749	Sequence 54749, A

C	19	36.4	1.9	1713	8	US-10-750-623-54749	Sequence 54749, A
C	20	36.2	1.9	537	6	US-09-925-065A-155337	Sequence 155337,
C	21	36.2	1.9	539	10	US-10-301-480-248761	Sequence 248761,
C	22	36.2	1.9	559	10	US-10-301-480-862170	Sequence 862170,
C	23	36.2	1.9	613	6	US-09-925-065A-506344	Sequence 506344,
C	24	36.2	1.9	4440	14	US-11-136-527-553	Sequence 553, App
C	25	35.8	1.9	608	6	US-09-925-065A-43711	Sequence 43711, A
C	26	35.8	1.9	608	9	US-10-301-480-144949	Sequence 144949,
C	27	35.8	1.9	608	10	US-10-301-480-758358	Sequence 758358,
C	28	35.8	1.9	616	6	US-09-925-065A-520203	Sequence 520203,
C	29	35.8	1.9	616	6	US-09-925-065A-520204	Sequence 520204,
C	30	35.6	1.9	576	6	US-09-925-065A-665158	Sequence 665158,
C	31	35.6	1.9	1437	14	US-11-098-686-8801	Sequence 8801, App
C	32	35.6	1.9	1457619	14	US-11-098-686-8739	Sequence 8739, App
C	33	35.4	1.8	581	6	US-09-925-065A-65665	Sequence 65665, A
C	34	35.4	1.8	581	9	US-10-301-480-166904	Sequence 166904,
C	35	35.4	1.8	581	10	US-10-301-480-760313	Sequence 760313,
C	36	35.4	1.8	712	6	US-09-925-065A-86811	Sequence 86811, A
C	37	35.4	1.8	712	9	US-10-301-480-188051	Sequence 188051,
C	38	35.4	1.8	712	10	US-10-301-480-801460	Sequence 801460,
C	39	35.2	1.8	484	6	US-09-925-065A-664128	Sequence 664128,
C	40	35.2	1.8	576	6	US-09-925-065A-665157	Sequence 665157,
C	41	35.2	1.8	585	6	US-09-925-065A-510680	Sequence 510680,
C	42	35.2	1.8	602	6	US-09-925-065A-568522	Sequence 568522,
C	43	35.2	1.8	629	6	US-09-925-065A-873796	Sequence 873796,
C	44	35.2	1.8	4834	14	US-11-000-688-1005	Sequence 1005, App
C	45	35.2	1.8	128978	8	US-10-775-169-345	Sequence 345, App

ALIGNMENTS

RESULT 1
US-09-925-065A-3789/c
Sequence 3789, Application US/0925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3789
LENGTH: 604
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-3789
Query Match 2.1%; Score 40.2; DB 6; Length 604;
Best Local Similarity 50.3%; Pred. No. 0.55;
Matches 96; Conservative 1; Mismatches 94; Indels 0; Gaps 0;
QY 176 CCGCTTTTGTGTAATGGAATGCAAACTCCCAAGTTTGCAGAAAGCTTG 235
DB 198 CTCCTGACTGACGCAATCTGCCACTTCGCTCCCAAGTGTGGATTACAGCGG 139
QY 236 AAACGTCGCGCAATATATCACTCATATTATCCATTTCATGACCAATTTTGAAG 295
DB 138 AACCACTGACCAACCAACCAAACTTTTATTTATATCAAGTACTTGAATTTAATGCC 79


```
RESULT 5
US-09-925-065A-562287/c
/ Sequence 562287, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 562287
/ LENGTH: 417
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-562287

Query Match
Best Local Similarity 2.1%; Score 39.6; DB 6; Length 417;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 243 CCCGGCAATTAATACTCCATATTATCCAAATTCAGACCAATTTTGAAGAAACAA 302
Db 230 CCCACATATAATAATAATACATATATAATAATAATAATAATAATAATAACAA 171
QY 303 ACAGCAGCTGGAATAAACAAGGGGCA 328
Db 170 ACAGCAGATGTAAATGACAGAGCA 145

RESULT 6
US-09-925-065A-124623/c
/ Sequence 124623, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 124623
/ LENGTH: 514
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-124623

Query Match
2.1%; Score 39.6; DB 6; Length 514;
```

```
Best Local Similarity 66.3%; Pred. No. 0.76;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 243 CCCGGCAATTAATACTCCATATTATCCAAATTCAGACCAATTTTGAAGAAACAA 302
Db 422 CCCACATATAATAATAATACATATATAATAATAATAATAATAATAATAACAA 363
QY 303 ACAGCAGCTGGAATAAACAAGGGGCA 328
Db 362 ACAGCAGATGTAAATGACAGAGCA 337

RESULT 7
US-10-301-480-75030/c
/ Sequence 75030, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ TITLE OF INVENTION: In the Human Genome
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 75030
/ LENGTH: 527
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-75030

Query Match
Best Local Similarity 2.1%; Score 39.6; DB 9; Length 527;
Matches 57; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 243 CCCGGCAATTAATACTCCATATTATCCAAATTCAGACCAATTTTGAAGAAACAA 302
Db 434 CCCACATATAATAATAATACATATATAATAATAATAATAATAATAATAACAA 375
QY 303 ACAGCAGCTGGAATAAACAAGGGGCA 328
Db 374 ACAGCAGATGTAAATGACAGAGCA 349

RESULT 8
US-10-301-480-688439/c
/ Sequence 688439, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ TITLE OF INVENTION: In the Human Genome
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 688439
/ LENGTH: 527
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-688439

Query Match
2.1%; Score 39.6; DB 10; Length 527;
Best Local Similarity 66.3%; Pred. No. 0.77;
```


NUMBER OF SEQ IN NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 841440
LENGTH: 656
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-841440

Query Match
Best Local Similarity 57.1%; Score 37.4; DB 6; Length 656;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATAGATCCCGAGCTTTTCACAAAAGGGGATCAGAGCTTGTAAAGCA 495
DB 132 CCGAAGACATTAATCTTGTCAAACTTAATGAATCAACAGACAGATTAAATGGA 73
QY 496 GTGGCGCTTGATACAAAGCAAAAATTAAGAGTGCAATTAAGAGCATCGAACAAAT 554
DB 72 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCAATGCACTCACTAAT 14

RESULT 13
US-09-925-065A-31269/c
Sequence 31269, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31269
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-31269

Query Match
Best Local Similarity 57.1%; Score 37.4; DB 6; Length 1766;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATAGATCCCGAGCTTTTCACAAAAGGGGATCAGAGCTTGTAAAGCA 495
DB 178 CCGAAGACATTAATCTTGTCAAACTTAATGAATCAACAGACAGATTAAATGGA 119
QY 496 GTGGCGCTTGATACAAAGCAAAAATTAAGAGTGCAATTAAGAGCATCGAACAAAT 554
DB 118 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCAATGCACTCACTAAT 60

RESULT 14
US-10-301-480-132507/c
Sequence 132507, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480

CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 132507
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-132507

Query Match
Best Local Similarity 57.1%; Score 37.4; DB 9; Length 1766;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATAGATCCCGAGCTTTTCACAAAAGGGGATCAGAGCTTGTAAAGCA 495
DB 178 CCGAAGACATTAATCTTGTCAAACTTAATGAATCAACAGACAGATTAAATGGA 119
QY 496 GTGGCGCTTGATACAAAGCAAAAATTAAGAGTGCAATTAAGAGCATCGAACAAAT 554
DB 118 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCAATGCACTCACTAAT 60

RESULT 15
US-10-301-480-745916/c
Sequence 745916, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 745916
LENGTH: 1766
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-745916

Query Match
Best Local Similarity 57.1%; Score 37.4; DB 10; Length 1766;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 436 CCGATATACAGATAGATCCCGAGCTTTTCACAAAAGGGGATCAGAGCTTGTAAAGCA 495
DB 178 CCGAAGACATTAATCTTGTCAAACTTAATGAATCAACAGACAGATTAAATGGA 119
QY 496 GTGGCGCTTGATACAAAGCAAAAATTAAGAGTGCAATTAAGAGCATCGAACAAAT 554
DB 118 GTGCAGCTTGACAAATTAATAAGAAAAAGAAAAAGTAAATCAATGCACTCACTAAT 60

Search completed: April 8, 2006, 03:55:08
Job time : 898.7 secs

This Page Blank (uspto)